The Quarterly Journal of Experimental Psychology

ISSN: 1747-0218 (Print) 1747-0226 (Online) Journal homepage: http://www.tandfonline.com/loi/pqje20

When comprehension elicits incomprehension: Deterioration of semantic categorisation in the absence of stimulus repetition

Upasana Nathaniel, Hannah E. Thompson, Emma Davies, Dominic Arnold, Glyn Hallam, Sara Stampacchia, Jonathan Smallwood & Elizabeth Jefferies

To cite this article: Upasana Nathaniel, Hannah E. Thompson, Emma Davies, Dominic Arnold, Glyn Hallam, Sara Stampacchia, Jonathan Smallwood & Elizabeth Jefferies (2017): When comprehension elicits incomprehension: Deterioration of semantic categorisation in the absence of stimulus repetition, The Quarterly Journal of Experimental Psychology, DOI: 10.1080/17470218.2017.1363793

To link to this article: http://dx.doi.org/10.1080/17470218.2017.1363793

Publisher: Taylor & Francis & The Experimental Psychology Society

Journal: *The Quarterly Journal of Experimental Psychology*

DOI: 10.1080/17470218.2017.1363793

When comprehension elicits incomprehension: Deterioration of semantic categorisation in the absence of stimulus repetition

Upasana NATHANIEL[1], Hannah E. THOMPSON[2], Emma DAVIES[1], Dominic ARNOLD[1], Glyn HALLAM[3], Sara STAMPACCHIA[1], Jonathan SMALLWOOD[1], Elizabeth JEFFERIES*[1]

1 Department of Psychology and York Neuroimaging Centre, University of York, UK

2 School of Psychology, University of Surrey, Guildford, UK

3 Department of Psychology, School of Human and Health Sciences, University of Huddersfield, UK

* Correspondence to:
Beth Jefferies,
Department of Psychology,
University of York,
York,
YO10 5DD, UK
Tel [+44] 01904 324368
Fax: [+44] 01904 323181
mailto: beth.jefferies@york.ac.uk

Running head: Semantic category fatigue

Keywords: fatigue, semantic, multimodal, refractory, satiation

Acknowledgements: The patient work was supported by a grant from the Stroke Association (TSA/12/02). EJ was supported by the European Research Council (SEMBIND – 283530). HET was supported by the Wellcome Trust [ref: 105624] through the Centre for Chronic Diseases and Disorders (C2D2) at the University of York. JS was supported by a grant from the European Research Council (Wandering Minds – 303701). We would like to thank Woody Chau for his help with data collection.

Abstract

Repetition improves retrieval from memory; however, under some circumstances, it can also impair performance. Separate literatures have investigated this phenomenon, including studies showing subjective loss of meaning following "semantic satiation", slowed naming and categorisation when semantically-related items are repeated, and semantic "access deficits" in aphasia. Such effects have been variously explained in terms of habituation of repeatedly-accessed representations, increased interference from strongly activated competitors, and longer-term weight changes reflecting the suppression of non-targets on earlier trials (i.e., retrieval-induced forgetting). While studies of semantic satiation involve massed repetition of individual items, competition and weight changes at the conceptual level should elicit declining comprehension for non-repeated items: this pattern has been demonstrated for picture naming but effects in categorisation are less clear. We developed a paced serial semantic task (PSST), in which participants identified category members amongst distracters. Performance in healthy young adults deteriorated with ongoing retrieval for non-repeated words belonging to functional categories (e.g., PICNIC), taxonomic categories (e.g., ANIMAL) and feature-based categories (e.g., colour RED – "tomato", "post box"). This decline was greatest at fast presentation speeds (when there was less time to overcome competition/inhibition), and for strongly-associated targets (which may have accrued more inhibition to facilitate earlier target categorisation). Deteriorating performance was also seen across words and pictures, consistent with a conceptual locus. We observed a release from deteriorating categorisation following a switch to a new category, demonstrating that this was not a general effect of time on task. Patients with semantic aphasia, who have deficient semantic control, maintained their performance throughout the categories, unlike younger adults: this finding is hard to reconcile with accounts of declining performance that propose a build-up of competition, since the patients should have had greater difficulty resolving such competition. These results instead suggest that declining performance on our goal-driven categorisation task was linked to the use of a controlled retrieval strategy by healthy young adults. Patients may not have inhibited related non-target knowledge to facilitate initial categorisation like younger volunteers, and consequently they were less vulnerable to declining comprehension in this paradigm. Together, these results demonstrate circumstances which produce declines in continuous categorisation in healthy adults.

Introduction

Repetition and priming largely have beneficial effects: they facilitate processing efficiency (Wagner, Desmond, Demb, Glover, & Gabrieli, 1997) and increase the accessibility of memory representations (Radeau, Besson, Fonteneau, & Castro, 1998). A similar benefit of repetition occurs for semantically-related items, where DOG primes a related word such as CAT: such effects are often explained in terms of automatic spreading activation between associated concepts (Badre & Wagner, 2002; Neely, 1977a). Nevertheless, several largely separate literatures have reported *declining* comprehension and semantic access following massed repetition of semantically-related sets: (i) patients with semantic "access" deficits show declining comprehension when small sets of semantically-related items are presented repeatedly; (ii) items are reported to 'lose their meaning' in massed repetition studies in healthy participants; a phenomenon referred to as 'semantic satiation'; (iii) psycholinguistic studies of healthy volunteers show poorer performance when semantically-related items are repeated: these effects are largely seen in picture naming, but can also affect comprehension (Campanella & Shallice, 2011; Harvey & Schnur, 2016; Wei & Schnur, 2015). In all of these separate literatures, there is a long-running debate about whether the effects arise at the lexical level (Crutch & Warrington, 2003; Damian, Vigliocco, & Levelt, 2001), within lexical-semantic links, or at the conceptual level (Belke, 2013; Gardner et al., 2012; Wei & Schnur, 2015). The underlying mechanisms are also unclear, with (a) some accounts proposing competition between currently-activated representations (Belke, Meyer, & Damian, 2005; Oppenheim, Dell, & Schwartz, 2010; Schnur, Schwartz, Brecher, & Hodgson, 2006), (b) other researchers noting that the effects are long-lasting, and are therefore more likely to reflect weight changes between associated items (Howard, Nickels, Coltheart, & Cole-Virtue, 2006; Oppenheim et al., 2010), or habituation of conceptual or lexical representations; and (c) patient studies emphasising that these effects are amplified by damage to executive processes outside the language/conceptual domain (Jefferies, Baker, Doran, & Ralph, 2007; Schnur et al., 2006). Finally, while many of these studies have involved the repeated presentation of individual items, declining comprehension should occur for non-repeated items if the effects arise at the conceptual level: research has already comprehensively demonstrated this pattern for picture naming (Belke, 2013; Harvey & Schnur, 2016; Howard et al., 2006) and here we show parallel effects in comprehension (see also Wei & Schnur, 2015).

Patients with semantic access impairment show "refractory effects", or declining comprehension in cyclical word-picture matching tasks (Jefferies et al., 2007; Warrington & McCarthy, 1983). When sets of semantically-related items are repeatedly presented, such that the target on one trial becomes a distractor on the next, patients become increasingly unable to select the target (Humphreys, 1997; Warrington & Cipolotti, 1996; Warrington & Crutch, 2004). This is only observed when the interval between one stimulus and the next is short (Campanella, Mondani, Skrap, & Shallice, 2009; Jefferies et al., 2007) and when the items are highly related in meaning (Cipolotti & Warrington, 1996; Crutch & Warrington, 2008; Forde & Humphreys, 1995; Forde & Humphreys, 1997; Jefferies et al., 2007). These effects have largely been documented in verbal comprehension tasks – i.e., word-picture matching (Cipolotti & Warrington, 1996; Jefferies et al., 2007; Schnur et al., 2006; Warrington & McCarthy, 1983; Warrington & Crutch, 2004). However, they have also been demonstrated in non-verbal judgements such as picture-picture matching (Forde & Humphreys, 2007; Gardner et al., 2012; Thompson, Robson, Lambon Ralph, & Jefferies, 2015), suggesting semantic access deficits can occur at a conceptual level. The mechanism that underpins this phenomenon is somewhat unclear (Mirman, 2011): declining performance on cyclical word-picture matching tasks has been linked to difficulty overcoming post-retrieval inhibition of selected representations (Gotts & Plaut, 2002) or to strong competition when several potential responses are activated (Forde & Humphreys, 1997; Jefferies et al., 2007; Schnur et al., 2006). The effect is strongest in patients with damage to prefrontal cortex (Campanella et al., 2009; Gardner et al., 2012; Schnur et al., 2006; Thompson et al., 2015), suggesting that it may reflect damage to control mechanisms that are necessary to maintain performance in the presence of strong competition on later trials and/or to overcome post-retrieval inhibition (i.e., when targets have to be re-selected having being inhibited on previous trials). In line with this view, patients with semantic aphasia (SA) show deficient semantic control across verbal and non-verbal tasks, characterised by difficulty supressing irrelevant aspects of knowledge and comprehending distant relationships and ambiguous meanings (Corbett, Jefferies, & Ralph, 2011; Jefferies & Lambon Ralph, 2006; Noonan, Jefferies, Corbett, & Lambon Ralph, 2009). These deficits give rise to declining comprehension in both cyclical word-picture matching and picture-picture matching (Gardner et al., 2012; Thompson et al., 2015).

Healthy participants typically do not show declining performance on cyclical word-picture or picture-picture matching tasks (Damian et al., 2001; Riley, McMahon, & de Zubicaray, 2015; some exceptions discussed below), although they do show increasing latencies in cyclical picture naming studies when sets of semantically-related items are named repeatedly (Kroll & Stewart, 1994; Vitkovitch & Humphreys, 1991). In some previous studies, these effects were explained in terms of competition from activated conceptual representations at the point of lexical selection: items drawn from the same semantic category activate each other via their shared conceptual features, and this might hinder retrieval of a specific object name because other activated concepts act as competitors (Belke, 2008; Belke et al., 2005; Damian et al., 2001). “Semantic blocking” effects in naming are sensitive to semantic variables, such as the strength of the association between the items (Abdel Rahman & Melinger, 2011), supporting the view that they reflect processes at the conceptual level.

A continuous picture naming paradigm can also elicit declining performance when conceptually-related items are presented (i.e., two animals) without the repetition of individual items (Belke & Stielow, 2013; Belke, 2013; Howard et al., 2006; Kleinman, 2013; Navarrete, Mahon, & Caramazza, 2010; Oppenheim et al., 2010; Runnqvist, Strijkers, Alario, & Costa, 2012; Schnur, 2014). For example, in the sequence GOAT, CAR, TOMATO, TRUCK, HORSE, the naming time for HORSE is slower than for GOAT due to their conceptual overlap, and this effect is found even at long 'lags' with many unrelated intervening items. Since this effect appears to be relatively long-term (cf. Wheeldon & Monsell, 1994), it is argued to reflect cumulative weight changes: each time an item from the animal category is named, the associative link between the word and animal features in the semantic system might be strengthened, increasing competition on future trials (Howard et al., 2006). Oppenheim et al. (2010) proposed that this type of incremental learning might not only reinforce the connections between semantic and lexical representations of targets, but also weaken semantic-lexical links for non-targets. Thus, semantic interference effects in naming can arise as a direct consequence of retrieval that renders related items less accessible (see also Anderson, Bjork, Bjork, & Jordan, 2000, for a related phenomenon in memory, termed "retrieval-induced forgetting").

Several recent studies used a continuous paradigm without item repetition to examine categorisation as opposed to picture naming, and observed cumulative *facilitation* rather than inhibition (Belke, 2013; Riley et al., 2015). Nevertheless, under some circumstances, healthy controls can show *declining* categorisation with repetition (Harvey & Schnur, 2015), and thus resemble patients with semantic access impairment. This pattern was observed in cyclical matching to a deadline when there were repeated presentations of the same target plus minimal delays between trials (Fabio Campanella & Shallice, 2011): these circumstances potentially create competition between the current target and previous targets (which have become distractors), with little time to resolve this competition or to recover from previous processing. In addition, Wei and Schnur (2015) reported semantic interference in a picture matching task, when the same response options were repeatedly used to probe associations with either related or unrelated concepts; in this study, there was initial facilitation (when semantically-related items were repeated at short lags; perhaps reflecting faster visual recognition for the probe when immediately following a related item), followed by longer-lasting inhibition (when related trials occurred at longer lags, perhaps reflecting response interference when a similar probe had led to a different decision on a previous trial).

Long-lasting declines in comprehension with repetition have also been reported in a third set of studies on "semantic satiation": this research reports that prolonged inspection and repetition of words results in a subjective loss of meaning (Jakobovits & Lambert, 1962; Smith & Klein, 1990). Semantic judgements are slowed under these circumstances but there is little effect on tasks such as lexical decision, suggesting that this effect again reflects effects at a semantic level (Cohene, Smith, & Klein, 1978; Neely, 1977b; Smith, 1984). Repetition is thought to cause temporary blocking of access to conceptual information (Frenck-Mestre, Besson, & Pynte, 1997; Pynte, 1991), potentially reflecting effects akin to neural fatigue or adaptation (Jakobovits & Lambert, 1962; Lambert & Jakobovits, 1960; Smith & Klein, 1990). Explanations for semantic satiation effects are similar to those above in that they anticipate spreading activation or inhibition to related conceptual representations; conceptual representations become "habituated" via repeated exposure, disrupting subsequent category judgements.

It is clear from the discussion above that semantic similarity across successive targets can produce both behavioural facilitation and inhibition on later trials, yet the situations that elicit these opposing effects remain largely unknown. Picture naming tends to show inhibition, whether or not items are repeated (e.g., Howard et al., 2006), while comprehension tasks tend to show facilitation (Belke, 2013; Riley et al., 2015) – although they sometimes show inhibition with item repetition (Campanella & Shallice, 2011; Wei & Schnur et al., 2015). There might be multiple mechanisms underpinning these diverse changes in performance, including priming, stimulus adaptation, competitive effects relating to spreading activation and lateral inhibition (thought to occur when target selection gives rise to the suppression of semantically-related competitors, that later become targets). The balance between these effects is likely to depend on the specific task that is presented. Some researchers have suggested that reductions in performance are restricted to paradigms requiring lexical selection (e.g., Damian et al., 2001; Riley et al., 2015); however, such effects can have a conceptual locus (e.g., Wei & Schnur, 2015) and therefore other task differences might be critical. A second difference between naming and word-picture matching is that naming involves generating representations from memory, while matching tasks present all the critical information as part of the paradigm. For this reason, retrieval demands are thought to be higher in picture naming. Controlled retrieval processes might give rise to declining performance on picture naming if participants retrieve targets by suppressing semantic neighbours, which later become targets (Oppenheim et al., 2010). These effects might not be prominent in most comprehension tasks, when the need for controlled retrieval is reduced.

Given these considerations, the current study used a paced serial semantic task (PSST) in which decisions to inputs were taken on the basis of a previously-encoded goal. Participants monitored a stream of inputs and pressed a button every time they detected a target that matched a particular category specified at the start of each block. We considered the ability of participants to sustain semantic processing over time: both *within* categories – by examining whether comprehension deteriorated over the course of each category as more related targets were presented, and *across* categories – by quantifying changes in performance across the experiment, as participants became generally fatigued. Neither targets nor distracters were repeated – therefore, the task resembled a continuous naming paradigm but required the comprehension and categorisation of meaningful inputs, rather than the production of speech. Although previous continuous categorisation experiments have shown facilitation (Belke, 2013; Riley et al., 2015) and, to our knowledge, there are no previous demonstrations of declining comprehension in the absence of stimulus repetition, our task might elicit cumulative inhibitory effects since semantic retrieval was guided in a top-down fashion (i.e., participants searched for an overlap between each input and the current goal category). This type of task might be expected to increase the need for controlled retrieval to each stimulus, as participants were required to focus on only currently-relevant aspects of knowledge. This aspect of the PSST might encourage the suppression of currently-irrelevant yet related concepts, which later become targets. The PSST paradigm allowed us to manipulate factors linked to both semantic representations themselves – such as strength of association, which should influence the spread of activation/inhibition between related concepts – and factors linked to control processes thought to play critical role in focussing retrieval on currently-relevant knowledge in the face of strong distractors or weak targets (Badre et al., 2005; Thompson-Schill, 1997; Whitney et al., 2011) – such as conditions of divided vs. undivided attention. In this way, the task maps onto contemporary accounts of semantic processing, which envisage conceptual representations that interact with control processes to support context- and task-appropriate semantic retrieval (Lambon Ralph, Jefferies, Patterson, & Rogers, 2016).

In the first part of the study, we manipulated factors that might influence the extent to which comprehension declines in healthy young adults. In Experiment 1, we investigated speed of presentation and semantic relatedness (both factors that produce declining comprehension in patients with semantic access deficits). We replicated the declining performance that we observed for thematic categories using taxonomic and specific feature judgements (Experiments 2 and 3), and found the same pattern of declining semantic performance with ongoing categorisation using a two alternative-forced-choice paradigm (Experiment 4). We also examined whether the effect extended beyond verbal comprehension to a task involving interleaved word and picture stimuli, to establish if this effect has a conceptual locus (which should transfer between modalities) (cf. Thompson et al., 2015; Wei & Schnur, 2015). Finally, we investigated whether the within-category decline in categorisation was increased by the requirement to divide attention in a dual task study, as would be expected if executive mechanisms guide retrieval in the face of strengthening competition or inhibition from semantically-related items (Experiment 5). Changes in the accessibility of meanings with ongoing categorisation in healthy young adults, in the absence of massed repetition, would have important implications for every-day comprehension.

In the second part of the study we considered how semantic aphasia (SA) influences performance on the PSST paradigm. These patients have well-documented deficits of semantic control alongside damage to left inferior frontal gyrus, which is implicated in the control of semantic retrieval by neuroimaging (Badre et al., 2005; Noonan, Jefferies, Visser, & Ralph, 2013) and neurostimulation studies (Whitney et al., 2011). Neuropsychology provided a straightforward way of distinguishing between alternative accounts of the within-category decline in categorisation: if these effects reflect a build-up of competition as more and more related items are presented, SA patients should have difficulty resolving this competition and therefore show an increase in the within-category decline in categorisation. In contrast, if retrieval-dependent declines in comprehension are a consequence of successfully dealing with competition on earlier trials through the inhibition of non-target information (cf. Anderson, Bjork, Bjork, & Jordan, 2000), SA patients might be expected to show attenuated or absent within-category decline, since their retrieval is relatively uncontrolled (e.g., Crutch & Warrington, 2005).

Experiment 1: Effect of speed of presentation and strength of association on thematic categorisation in healthy participants

Rationale

This experiment tests whether semantic retrieval declines over time (even in the absence of explicit item repetition), and if there is a release from this effect when the task switches to a new category. We considered whether this pattern of declining categorisation would be greater at a fast rate of presentation. Previous research has found declining performance with both fast and slower presentation speeds: fast speeds might allow a greater build-up of competition from previously-presented semantically-related items (Fabio Campanella & Shallice, 2011), yet satiation effects occur when inputs are presented for long durations (Smith, 1984) and performance in continuous paradigms declines even over relatively long lags, reflecting slower cumulative weight changes (Oppenheim et al., 2010). We therefore compared performance at fast (1.1s) and slower speeds of presentation (2s) to establish whether this manipulation would alter the extent to which semantic categorisation declined in the PSST paradigm. In addition, we tested whether the effect would vary as a function of the strength of association between the presented words and the target category. Larger effects of within-category decline might occur for weakly-associated words if declining comprehension is caused by a build-up in competition, since weak associations should be harder to retrieve in the face of competition from previously-presented strong category members. Alternatively, the effect could be larger for strongly-associated words, if it reflects weight changes that render related non-target items on previous trials less accessible (Oppenheim et al., 2010). The literature on semantic satiation shows stronger detrimental effects of repetition for strong vs. weak associates (Balota & Black, 1997). Similarly, patients with semantic access deficits show declining cyclical word-picture matching for sets of closely related items – but not for sets of repeated but unrelated or distantly-related items (Crutch & Warrington, 2003). Warrington and Cipolotti (1996) found a detrimental effect of cycle in these patients even when items in the last cycle were replaced with new items from the same category, suggesting that spreading activation between strongly-related concepts is the likely cause of this decline. Therefore decline in performance was compared for strong and weak members of a thematic category (e.g., PICNIC – strong category member = "sandwich", weak category member = "wasp").

Method

Participants: 24 undergraduate students (16 females and 8 males) from the University of York participated in the experiment in exchange for course credit or a payment of £5. The mean age of the students was 19 years (range of 18-24). All participants were native English speakers. Ethical approval was obtained from the Department of Psychology Ethics Committee at the University of York. All participants gave written informed consent.

Task and Design: The 'Paced Serial Semantic Task' or PSST required rapid semantic association judgements that linked spoken words to a thematic category, such as PICNIC or HOSPITAL. Participants were asked to classify spoken words in terms of whether they were associated with the target categories. Two factors were manipulated in a repeated-measures design: (1) the strength of association between the target and category (strong or weak), and (2) presentation speed (fast: 1.1s or slow: 2s). The experiment additionally looked at (3) effects of 'within-category decline' (comparison of task performance in the first half compared with the second half of each category), and (4) 'across-category decline' (i.e., decline in performance across the testing session).

Materials: Twenty different category labels were used (such as PICNIC) with 60 items in each category. 20 items were related to the category, including 10 targets that were strongly related to the category label, such as "sandwich"; and 10 that were distantly related, such as "wasp", while the remaining 40 items were unrelated to the category (e.g. "exam") – these were recycled items from other categories (see Appendix A in the Supplemental Material for a complete list of items used). Target words were selected using the Edinburgh Associative Thesaurus (EAT; Kiss, Armstrong, & Milroy, 1973), supplemented by a pilot study in which ratings were collected for the relatedness of each word to the category label. Participants (N = 16) used a 7-point Likert scale to judge relatedness, and items were categorised as strongly related (> 5.5), weakly related (2.2 - 5.5) or unrelated (< 2.2).

Procedure: The experiment was presented using E-Prime 2.0 (Psychology Software Tools, Sharpsburg, PA). Category names were presented as written words that remained visible throughout the block, to reduce demands on working memory. There was an equal distribution of strong and weak targets in the first and second half of each category. The order of categories and items was fully counterbalanced between subjects (there were parallel versions of the experiment utilising two presentation orders; each of these was presented to half of the participants, such that effects at the group level could not reflect effects specific to one order of presentation). These details were repeated across all experiments below.

Participants were asked to press '1' each time they heard a word that was related to the category, and not to press for unrelated words. Thus the task required sustained and rapid attention to semantic information. Each participant was presented with all 20 categories, 10 at one speed (e.g., with a 1.1 second gap between each auditory word) and 10 at another speed (e.g., 2 second ISI), with the two speeds counterbalanced using an ABBA or BAAB design.

Results

The main dependent measure in all experiments was response sensitivity (*d'*), which accounts for response bias (the general tendency to respond *yes* or *no*; Stanislaw & Todorov, 1999). A higher *d'* score reflects better response sensitivity (i.e., the ability to correctly recognise targets and reject distractors – as opposed to making 'false alarms' on non-target items). As there was an equal distribution of strong and weak targets in the first half and second half of each category, within-category changes in performance were examined by computing *d'* separately for these two halves of each category. In Experiment 1 and subsequent experiments, generalised linear models (GLMs, using generalised estimating equations) were used to analyse response sensitivity for each category and for each participant, including within-subject fixed-effects of within-category position (first vs. second half of each category), across-category fatigue (first vs. second half of the entire experiment), speed of presentation (1.1 vs. 2s), and semantic relatedness (*d'* scores computed separately for targets that were strongly or weakly related to the category), in a fully-factorial model that included all interaction terms for these predictor variables. These *d'* scores were therefore computed across sets of trials, retaining information about performance per category per participant (i.e., categories were treated in the same way as individual trials in a classic 'by-items' analysis). The GLMs allowed for random variation in the intercept across participants. Average RT was entered for each of the first and second half of each category as a covariate (i.e., the average RT for correct responses per condition and participant) in this and all subsequent GLMs. Given that the task required participants to respond before a deadline (i.e., before the onset of the next item, rather than as quickly as possible), RT was not expected to be sensitive to the effects of interest but by including average RT as a covariate in the analysis, changes in RT were accounted for over the course of a block of trials. This would allow focus on response sensitivity while simultaneously accounting for the possibility of a response accuracy trade–off.

Response sensitivity in each condition is shown in Figure 1. Table 1 shows the results of the GLM analysis (alongside conventional repeated-measures ANOVA of response sensitivity, which revealed the same effects of the experimental factors). There was a main effect of speed, indicating better performance at a slower rate of presentation. There was also a main effect of association strength: sensitivity was lower for weakly related items in comparison to stronger associations. There was no significant main effect of within-category position or across-category fatigue ($p > .1$). However, there was a significant interaction of relatedness and within-category decline: participants made more errors in categorisation towards the end of each category especially for the strongly associated targets. Post-hoc tests with Bonferroni correction indicated this decline in sensitivity 'within' each category affected performance on strongly related targets: t (23) = 2.34, p = .028, but not weakly related targets: t (23) = .105, p = .917. There was also a significant interaction between speed and relatedness: participants found it harder to identify weak items at the faster speed in comparison to the slower speed. Bonferroni corrected t-tests indicated significant effects of relatedness at both speeds, with a larger effect at the fast speed: t (23) = 7.64, $p < .001$, in comparison to the slow speed: t (23) = 5.37, $p < .001$. All other interaction terms were non-significant. A more detailed breakdown of performance, showing hits, correct rejections, false alarms and misses, is reported in Table 1 in Supplemental Materials. This shows few false alarms; due to the paced nature of the paradigm, participants tended to fail to respond to targets when performance was poor.

-Figure 1 about here-

-Table 1 about here-

Summary of Experiment 1

Even though individual words were not repeated, as in typical semantic satiation or cyclical matching experiments, performance showed a cumulative decline across the targets in each category. Within-category decreases in response sensitivity were greater for targets strongly associated with the category label (e.g., PICNIC – "sandwich"), compared with weakly associated targets (e.g., PICNIC – "wasp"). Performance on the task was also influenced by speed of presentation: participants showed poorer performance at faster speeds. There was no decline in performance over the course of the experiment (i.e., across-category fatigue effects were not significant). Thus, the PSST revealed several of the hallmarks of "semantic access deficits" in healthy subjects – namely, declining performance with on-going semantic retrieval; greater effects when there was a deadline to respond (Fabio Campanella & Shallice, 2011); plus a more substantial decline for more strongly-related items, suggesting that this effect reflects the spread of activation or inhibition within semantic representations.

Additionally, in line with studies of continuous picture naming (which examine performance as a function of how many trials lie between presentations of semantically-related items), we examined accuracy as a function of the number of distracters since the last target (reported in Supplemental Materials). We found no evidence that the within-category fatigue effect was reduced in magnitude when the number of intervening items was greater (see Supplemental Figure 1). This suggests that ongoing neural activation from previous related trials was unlikely to be the cause of the decline in performance, since such an account would predict the greatest decline in performance at zero lag, when there has been no time for decay in activation. These results are more consistent with the view that longer-lasting weight changes underpin the effect (cf. Wei & Schnur, 2015).

Experiment 2: Taxonomic category matching in healthy participants

Rationale

This experiment provided a replication of the within-category decline effect in Experiment 1 using taxonomic categories as opposed to thematic categories. The target items shared common features (e.g., eyes and fur, for the category ANIMALS). If within-category decline reflects spreading activation within the semantic system that interferes with the categorisation of incoming items, this effect should be observed for targets with strong featural overlap.

Method

Participants: 24 undergraduate students (16 females and 8 males) were recruited from the University of York, and received course credit or a payment of £5 for their participation. The mean age of the students was 21 years (range of 18-30). All participants were native English speakers.

Task and procedure: Targets were strong members of each taxonomic category (e.g., “apple”, “orange”, “grapes” for FRUITS). Twenty different category labels were used (VEHICLES, FLOWERS, BIRDS, etc.). For each category, there were 20 related items (e.g., VEHICLES – “car”) and 40 unrelated items (e.g., VEHICLES – “meerkat”) in each category – the unrelated items were targets from other categories (see Appendix B in the Supplemental Material for a complete list of items used). Unlike Experiment 1, this experiment did not include manipulations of speed or relatedness. Items were presented at a speed of 1.1s. The experiment was presented using E-prime 2.0 (Psychology Software Tools, Sharpsburg, PA).

Results

Response sensitivity is shown in Figure 2. The effects of two fixed-effects were examined in a GLM: (1) ‘across-category fatigue’ (comparison of task performance in the first half of session compared with the second half of session); and (2) ‘within-category decline’ (comparison of task performance in the first compared with the second half of each category) in a fully factorial model, including RT as a covariate of no interest. A parallel analysis using repeated-measures ANOVA revealed the same pattern of results; see Table 2 for *Wald* χ^2, *F* and *p* values.

The GLM analysis found a significant main effect of within-category decline, indicating poorer performance towards the end of each category. There was little evidence that performance changed across the experiment (i.e., no across-category effect). There was also no significant interaction of within-category and across-category decline.

-Figure 2 about here-

-Table 2 about here-

Summary of Experiment 2

Taxonomic categorisation was easier overall than the thematic judgements used in Experiment 1 (as reflected in larger *d’* scores). However, a similar pattern of declining performance *within* each category was observed. Since performance did not decline *across* categories, this effect was did not reflect a general difficulty in sustaining attention to the task.

Experiment 3: Specific feature matching in healthy participants

Rationale

Experiment 3 investigated whether the within-category semantic decline observed in Experiments 1 and 2 would occur when items were categorised on the basis of single feature, such as colour (e.g., "post-box", "tomato" and "Santa" for the category RED). Feature matching is a demanding semantic task that requires executive resources to focus semantic retrieval on the feature relevant to the task, and away from dominant aspects of knowledge (Badre et al., 2005; Jefferies, 2013): targets in this experiment were not globally related to the category being probed, and shared few (if any) features, except for the feature specified in the instructions (e.g., the targets "pancakes", "blackboard", "postcard" share the feature FLAT but are not globally related). If the requirement to maintain a narrow focus of conceptual retrieval underpins the pattern of deteriorating categorisation, these effects would be expected to be maintained in this experiment. If, in contrast, strong global semantic relationships between target items are necessary for within-category decline in performance, this effect should be reduced in magnitude or even eliminated in this experiment.

Method

Participants**:** There were 24 participants (18 females and 6 males); recruited from the University of York, in exchange for course credit or payment of £5 Mean age of the students was 19 years (range of 18-24). All participants were native English speakers.

Task and procedure: The paradigm was similar to that in Experiment 2, except categorisation was based on a specific feature of the presented items. For example, participants were shown a category, such as RED and were asked to classify spoken words (such as, "tomato", "post-box", "Santa"), in terms of whether they matched this specific feature. Twenty-two different category labels were used (e.g., NOISY, FLAT, HOT, etc.) with 60 items in each category: 20 were related items (e.g., NOISY – "vacuum cleaner"), and 40 were unrelated items (e.g., NOISY – "caramel") taken from other categories (see Appendix C in the Supplemental Material for a complete list of items used). Each participant was presented with all 22 categories. The order of categories and items was fully counterbalanced between subjects.

Results

Results are shown in Figure 3. This experiment assessed the effects of two within-subjects factors in a GLM: (1) 'across-category decline' (comparison of task performance in the first half of the experiment compared with the second half); and (2) 'within-category decline' (comparison of task performance in the first compared with the second half of each category). These fixed effects were entered into a fully-factorial model. A parallel analysis using repeated-measures ANOVA was conducted on sensitivity, which yielded similar results (see Table 3 for *Wald* χ^2, *F* and *p* values).

There was a significant main effect of within-category decline: performance was better at the beginning than the end of each category. There was no main effect of declining performance across-categories ($p > .1$). There was a significant across-category by within-category interaction, indicating a greater decline in performance 'within' each category towards the end of the experiment (see Figure 3). This was supported by Bonferroni t-tests, which indicated a highly significant within-category decline for the second half of the session: t (23) = 4.40, $p < .001$, but not for the first half of the session ($p > .1$).

-Figure 3 about here-

-Table 3 about here-

Summary of Experiment 3

This experiment provided a second replication of the cumulative decline in categorisation performance within categories, in the absence of item repetition; however, in this case, the pattern was only apparent in the second half of the session. The feature-based classification task used in Experiment 3 was more demanding than the thematic and taxonomic categorisation tasks used in Experiments 1 and 2, and the within-category decrease in performance on this executively-demanding semantic task might have been maximal towards the end of testing session, when cognitive control was likely to be lower. Most importantly, this experiment shows that the within-category decline effect extends to situations in which there is not a strong global relationship between the targets. The effect might therefore not emerge from strengthening activation in sets of globally-related concepts, but might instead reflect interactions between semantic goal representations (i.e., targets are 'thin', or 'red', or 'round') and the conceptual store.

Experiment 4: Within-category decline in categorisation across modalities in healthy participants

Rationale

Experiment 4 considers whether the systematic decline in the meaning of an item occurs in a manner that is independent of a specific modality. Declining performance in the cyclical matching paradigm in patients with aphasia has largely been documented using verbal comprehension tasks – i.e., word-picture matching (Cipolotti & Warrington, 1996; Jefferies et al., 2007; Schnur et al., 2006;

Warrington & McCarthy, 1983). It has been suggested that this effect may be exclusive to auditory or verbal materials (Crutch & Warrington, 2008; Warrington & Crutch, 2004). Similarly, in healthy participants, declining performance on cyclical paradigms has been linked to lexical competition during speech production (rather than conceptual retrieval) (e.g., Belke et al., 2005; Harvey & Schnur, 2015; Howard et al., 2006), while in the satiation literature, it has been suggested that declining comprehension comes about due to adaptation of orthographic-to-semantic links (Tian & Huber, 2010) – consequently, these effects of repetition may be restricted to the verbal domain. Nevertheless, SA patients show declining performance across cycles for both word-picture and picture-picture matching tasks (Forde & Humphreys, 2007; Gardner et al., 2012), suggesting that semantic access deficits can occur at a conceptual level, and similar results were obtained recently in healthy participants (Wei & Schnur, 2015). This is consistent with the proposal that semantic cognition draws on heteromodal representations and control processes that operate across modalities. This study characterised the decline in performance for word targets, picture targets and an interleaved condition in which related items were presented as both words and pictures on different trials. If the decline in performance arises at a conceptual level, this effect should not be diminished for the interleaved condition. This pattern would allow ruling out accounts of within-category decline that involve fatigue/adaptation or competition within lexical-level representations.

In this experiment, the PSST paradigm was also modified: participants were asked to make a response on each trial (pressing one of two buttons to indicate if the item was a member of the category, or not), and there were equal numbers of targets and distractors. This two alternative-forced-choice (2AFC) design minimised the effects of response bias (relative to the paradigm used above, in which participants only responded when a target was present) and, most importantly, allowed us to characterise within-category changes in performance not only in terms of hits but also correct rejections, to examine if participants were updating their working definition of the category over the set of trials. If participants showed a similar within-category reduction in accuracy for both targets and non-targets, it would suggest a reduced ability to retrieve the relevant information; in contrast, if they only showed a change for targets, it could suggest a narrower definition of the category is being acquired as the category progresses (i.e., a shift in response criteria).

Method

Participants: 24 participants, native English speakers aged between 18-30 years old, were recruited from the University of York in exchange for course credit or a payment of £5

Task and Design: The target categories in this experiment were thematic, in line with Experiment 1. Within-category manipulations involved: (i) stimulus modality (auditory words, pictures and interleaved auditory words and pictures) and (ii) within-category position (first half of each category was compared to the second half). The possibility of declining performance across categories was not examined in this experiment, since performance would have been influenced by the order in which the three modality conditions were presented. Strength of association ratings were also used to split the verbal targets into strong and weak. The effect of this factor is reported in a cross-experiment comparison below.

Materials: There were 30 categories, 10 per modality (words, images, interleaved). Each participant saw each category only once (in one of the three modality conditions). There were 40 items in each category (20 related and 20 unrelated to the category). In the visual condition, the images were colour photographs on a white background. In the auditory condition, the word stimuli were audio files (see Appendix D in the Supplemental Material for a complete list of items used).

Procedure: The experiment was presented using E-Prime 2.0 (Psychology Software Tools, Sharpsburg, PA). Participants completed 15 practice trials (5 trials for each modality), before proceeding to the experimental trials. At the start of each category block, the category label was written on screen until the participant pressed the spacebar to continue. Pilot testing indicated that participants were unable to perform the interleaved condition at 1.1s, so a slightly slower presentation speed of 1.3s was adopted. Each item was presented for 1.3s and this was the deadline for responding. Participants pressed one of two buttons to indicate if the item was related or unrelated to the target category. The order of stimulus modality, categories, and items within each category was fully counterbalanced across participants (although in the interleaved condition, a spoken word was followed by a picture).

Results

Results are shown in Figure 4. The GLM included *modality* (words vs. pictures); *interleaving* (single modality vs. interleaved words/pictures), and *within-category decline,* as fixed within-subjects effects and controlled for RT as a covariate. Parallel analysis using repeated-measures ANOVA on response sensitivity obtained similar results (see Table 4 for *Wald* χ^2, *F* and *p* values).

There was a main effect of interleaving: lower performance for interleaved vs. non-interleaved trials. There were no other main effects ($p > .1$). There was an interaction between within-category decline and interleaving: the cumulative decrease in performance 'within' each category was larger for interleaved than non-interleaved trials. This was supported by post-hoc tests, which indicated a significant decline for the interleaved condition: t (21) = 2.17, p = .042, but not the non-interleaved trials (p > .1). There was also a significant interaction between modality and interleaving conditions, reflecting a greater effect of interleaving for pictures than words. Bonferroni corrected t-tests indicated a highly significant effect of interleaving in the pictures modality: t (21) = 3.20, p = .004, but not the words modality (p > .1). Other interactions were not significant (p > .1).

-Figure 4 about here-

-Table 4 about here-

Summary of Experiment 4

In an alternative design using 2AFC decisions, a within-category decline was observed for both words and pictures, demonstrating that this effect does not reflect changes within lexical-semantic representations. Instead, within-category decreases in categorisation appear to reflect processes at a conceptual level. The within-category decline effect was also more marked for interleaved blocks, containing both word and picture targets, relative to single-modality blocks, suggesting that the effect can accumulate across these inputs. Interleaved blocks were more difficult, presumably because of the greater need for attentional control and switching: this may explain why the magnitude of within-category decline was strongest in this condition. Whether within-category decline interacts with the availability of control resources is tested in Experiment 5. The current experiment also demonstrated a reduction in performance for both target and non-target trials (i.e., an increase in both misses and false alarms, see supplemental material). This result suggests that the within-category decline effect is not a change in participants' criteria for category membership. Instead it is more likely to occur because of an increasing inability to identify the targets and reject the non-targets.

Experiment 5: Effect of divided attention on within-category decline in healthy participants

Rationale

Research suggests that semantic cognition involves an interaction between conceptual representations and control processes that focus retrieval on currently-relevant aspects of knowledge (Jefferies, 2013; Lambon Ralph, Jefferies, Patterson, & Rogers, 2016). This type of control over retrieval may be partly achieved by domain-general executive mechanisms (although there might also be neurocognitive mechanisms that support semantic or memory control specifically; Davey et al., 2016; Noonan, Jefferies, Visser, & Ralph, 2013). Previous research has already shown that the requirement to perform a secondary task concurrently with semantic retrieval disrupts access to non-dominant knowledge (Almaghyuli, Thompson, Lambon Ralph, & Jefferies, 2012). Thus, if within-category decline reflects an increase in either competition or difficulty retrieving targets following longer-term inhibitory weight changes, control mechanisms that can resolve competition or promote weak but currently-relevant information should become more important towards the end of each category. Under these circumstances, the requirement to do two tasks at once might have a particularly detrimental effect on comprehension later in the categories.

The neuropsychological literature already points to the importance of control processes since patients with semantic access deficits (who show semantic blocking effects on cyclical word-picture matching tests) tend to have damage to left PFC and problems focussing retrieval on currently-relevant information that correlates with general executive dysfunction (Fabio Campanella et al., 2009; Gardner et al., 2012; Jefferies et al., 2007; Jefferies & Lambon Ralph, 2006; Schnur et al., 2009; Schnur et al., 2006; Thompson et al., 2015). There are also several findings from the previous experiments reported here that suggest a role for control processes in within-category decline in categorisation. The interaction of within- and across-category decline seen in Experiment 3 but not the other experiments could reflect the importance of executive resources for the difficult feature matching task, particularly at the end of each category (due to competition and/or weight changes). Within-category decline was also greater in the interleaved condition in Experiment 4, which required participants to switch between input modalities and may have had higher control demands. To directly test the importance of executive resources, Experiment 5 examined within-category decline in the paced serial semantic task with and without the requirement to perform a secondary task.

Method

Participants: 24 undergraduate students (20 females and 4 males) from the University of York took part, in exchange for course credit or a payment of £5. The mean age of the students was 20 years (range of 18-30). All participants were native English speakers.

Task Design: This experiment used a thematic category matching task (as for Experiment 1) and manipulated: (i) condition (single or dual), and (ii) strength of association (strongly vs. weakly associated targets). The effects of within-category and across-category decline were also examined. In the single task condition, participants were asked to press a button when they detected targets that related to the category (identical to Experiment 1). In the dual task condition, participants performed the same semantic task, except this time they were also asked to count triangles that appeared on the screen over the course of the category and report this number at the end.

Materials: All stimuli were taken from Experiment 1. Twenty thematic category labels were used (CHURCH, AIRPORT, MUSIC FESTIVAL, etc.) with 60 spoken items presented in each category. 20 items were related to the category, including 10 targets that were strongly related, (e.g., CHURCH – "priest"), and 10 that were distantly related to the category (e.g., CHURCH – "bread") while the remaining 40 items were unrelated to the category (e.g. CHURCH – "oyster") – these were recycled from other categories (see Appendix A in the Supplemental Material for a complete list of items used). 10 categories were combined with the secondary task while 10 were presented under single task conditions. In the dual task condition, patterned triangles were presented on the screen (with 30 triangles appearing overall, 15 triangles distributed in the first half and 15 triangles in the second half of the session). In order to minimise the difficulty of the dual condition, only 2 – 4 triangles were presented per category, and these appeared on randomly-selected trials.

Procedure: Each session began with three practice blocks – the first block involved categorising spoken words (i.e. single task condition); the second block involved presentation of triangles (without any auditory stimuli), participants were asked to count the triangles that appeared from time to time on a blank screen and write down the number they had seen; the third practice block combined the two tasks. There were three categories with 30 trials (5 strongly related, 5 weakly related and 20 unrelated items) in each of the practice blocks. After the practice blocks, participants were presented with 20 experimental categories, using an ABBA or BAAB design for the single and dual-task conditions. The order of conditions, categories and items was fully counterbalanced between subjects.

Results

The results of Experiment 5 are presented in Figure 5. GLM analysis in this experiment included *secondary task condition* (single vs. dual task), *relatedness*, *across-category* and *within-category* as fixed within-subjects effects and included RT as a covariate. Parallel analysis using repeated-measures ANOVA on the sensitivity data indicated similar results (see Table 5 for *Wald* χ^2, F and p values).

The analysis revealed significant main effects of secondary task condition (reduced sensitivity in the dual task compared to the single task), relatedness (lower sensitivity for the weak than strong targets) and a marginal main effect of across-category decline (sensitivity scores declined overall from the first half to the second half the session). There was no main effect of a within-category decline, but there was a significant interaction between relatedness and within-category performance: the decline in categorisation was greater for strong than weak items. This was supported by post-hoc tests with Bonferroni correction, that showed highly significant within-category decline for strong items: t (23) = 7.41, p < .001, and near-significant decline for weak items: t (23) = 1.99, p = .058.

There was a trend-level three-way interaction between dual task condition, relatedness and within-category decline. This was explored by analysing performance for the strong and weak targets separately. The dual task by within-category interaction was significant for strong items: t (23) = 2.83, p = .009, but not for the weak items (p > .1). Other interactions were not significant (p > .1).

-Figure 5 about here-

-Table 5 about here-

Summary of Experiment 5

The strong items showed a substantial effect of within-category decline, especially under dual task conditions; i.e., divided attention increased the decline in categorisation seen towards the end of each category. Thus, this experiment provides tentative support for the view that a reduction in executive control in healthy young volunteers increases the effects of within-category decline. Executive resources might allow participants to overcome competition from strongly-activated distractors, or the effects of retrieval-induced inhibition (i.e., the suppression of related items on previous trials).

Cross-experiment analysis of effects of within-category decline in healthy participants

Declining comprehension within and across categories

Given the similar structure of several of the experiments above, a meta-analysis was conducted examining the magnitude of both within-category decline (i.e., performance in the first and second half of each category) and across-category decline (i.e., performance in the first and second half of each experiment). All experiments employing a 1.1s presentation speed were included. This included the fast presentation condition of Experiment 1, Experiments 2 and 3, plus the single-task condition from Experiment 5. This analysis collapsed across strong and weak targets in Experiments 1 and 5. GLM analysis included *across-category* and *within-category* as within-subjects fixed effects. *Experiment* was included as a between-subjects factor to establish if the magnitude of within- or across-category decline varied across these experiments. RT per condition and participant was again used as a covariate (see Table 6 for results). There was a main effect of Experiment: the taxonomic categorisation task (Experiment 2) was easier than thematic matching

(Experiment 1), specific feature matching (Experiment 3) and the single-task condition in Experiment 5. There was a significant main effect of within-category decline across these four experiments (see Figure 6). However, there was no overall change in performance across experiments (i.e., no across-category decline) and no significant interactions ($p > .1$). A decline in performance was revealed towards the end of each category across all of these experiments.

-Figure 6 about here-

-Table 6 about here-

To further characterise the way in which performance changed across successive items within categories, the average number of hits and false alarms was computed for each individual item position across these experiments. Each category included 60 items (20 targets and 40 non-targets, which were presented in different positions across categories and participants). From this data, we computed the average percentage of hits and false alarms individually for each position (reported in Figure 7). The analysis revealed a largely continuous decline in hits, with no substantial increase in false alarms, consistent with the analyses presented above (see Figure 7). Item position correlated with the percentage of hits across the four experiments; this effect was greatest in Experiment 3 ($r = .636$, $p < .001$), followed by Experiments 1 and 5 ($r = .559$, $p < .001$; $r = 452$, $p < .001$ respectively), and Experiment 2 ($r = .291$, $p = .024$).

-Figure 7 about here-

Interactions with strength of association

In an additional analysis across experiments, the relationship between within-category decline and the strength of association between stimuli and target categories was examined. This analysis included data from Experiment 1 (strong vs. weak targets, collapsed across the two speeds – 2s and 1.1s), Experiment 4 (strong vs. weak word targets, collapsed across interleaved and non-interleaved conditions) and Experiment 5 (strong vs. weak targets, collapsed across the single and dual conditions). The picture condition from Experiment 4 was omitted since verbal measures of strength of association may not apply to picture-based decisions.

A GLM examining response sensitivity included *relatedness* and *within-category* as fixed effects within-subjects, *Experiment* as a between-subjects factor, and RT as a covariate (see Table 7 for results and parallel analysis using repeated-measures ANOVA). The analysis revealed a main effect of relatedness: overall sensitivity scores were lower for weak compared to strong targets (see Figure 8). There was a main effect of within-category decline: sensitivity declined towards the end of each category across the three experiments. Importantly, there was a significant interaction between relatedness and within-category decline. Post-hoc tests with Bonferroni correction indicated a significant within-category decline for strong targets: t (69) = 3.83, p < .001, but not weak associations (p > .1). Thus, performance declined more substantially for strongly than-weakly related targets across experiments.

There was also a significant interaction between task and relatedness: there was a stronger effect of relatedness in Experiment 5 (which involved divided attention) and in Experiment 4 (which involved interleaved presentation), in comparison to Experiment 1 (a simpler thematic matching task). Post-hoc tests with Bonferroni correction indicated a significant effect of relatedness in all experiments, which was largest in Experiment 5: t (23) = 21.38, p < .001, followed by Experiment 4: t (21) = 8.28, p < .001, and smallest in Experiment 1: t (23) = 6.41, p < .001. In this way, the effects of strength of association appeared to be greater in more executively-demanding paradigms.

-Figure 8 about here-

-Table 7 about here-

Experiment 6: Retrieval-induced changes in categorisation in patients with semantic aphasia

Rationale

The findings so far show that young adults find it progressively harder to rapidly categorise items drawn from a single category. This effect was greater for strong associations, suggesting it resulted from spreading activation or inhibition between related items. It was also increased by divided attention, suggesting that control may be employed to overcome the harmful effects of this spreading activation/inhibition. However, the data do not adequately discriminate between theoretical accounts which propose that declines in categorisation reflect (i) an increasing need for controlled retrieval later in each category – following, for example, the activation of a broad semantic field (giving rise to strengthening competition) and (ii) the unintended consequences of controlled retrieval earlier in the list (for example, if semantically-related but irrelevant information is inhibited on earlier trials, and then becomes relevant). By the first account, in a category such as PICNIC, many relevant items are activated by the end of the list (WASP, CAKE, SUNSHINE), making it difficult to focus retrieval on a specific item (e.g., RUG). By the second account, participants facilitate their performance on initial items by suppressing aspects of picnic knowledge not relevant to those items (e.g., for WASP – selectively focus on flapping away wasps; inhibit aspects of picnics that are relaxing). This makes it harder to see the relevance of later targets (e.g., relaxing on a RUG). The vulnerability of strongly-related items to deteriorating categorisation appears more consistent with the second account, since strong associations relate to central aspects of the category, which should be suppressed during the retrieval of weakly-associated items early in the list. In contrast, rising competition should impair the retrieval of weak associations to a greater extent. The secondary task effects in Experiment 5 are potentially consistent with either account, since control processes could compensate for strong competition, or for the earlier inhibition of knowledge. In order to select between these two alternative theoretical accounts, a final experiment examined aphasia patients with semantic deficits following stroke. This population has been previously shown to have deficits in control over semantic retrieval (Corbett et al., 2011; Jefferies & Lambon Ralph, 2006; Noonan et al., 2009). We tested the consequence of this deficit on the PSST paradigm to discriminate between the two accounts of within-category decline outlined above. (i) If within-category decline in comprehension reflects an increased need for controlled retrieval, for example, following increased competition from previously-activated semantically-related items, patients with semantic control deficits should show *greater* within-category decline relative to controls. (ii) If within-category decline reflects the consequences of employing control on earlier trials, to suppress semantic competitors, patients with semantic control deficits should show *decreased* effects, especially relative to healthy young volunteers.

Using the PSST paradigm we characterised changes in semantic cognition in terms of: (1) the effect of strength of association (strong vs. weakly-related targets); (2) within-category decline – e.g., contrasting the beginning and end of each category; (3) across-category decline – e.g., general cognitive fatigue which might produce deteriorating performance on the paradigm over the course of each testing session. The patients were expected to be disproportionately impaired at retrieving weak as opposed to strong associations, in line with previous findings (Corbett et al., 2011; Jefferies & Lambon Ralph, 2006; Noonan et al., 2009). A secondary research question concerned the extent to which the patients would show cognitive fatigue across the session on a task requiring sustained semantic attention.

Method

Participants

Patients: Twelve stroke aphasic patients (eight female, four male) were recruited from stroke clubs and speech and language therapy services in York and Leeds, UK. All patients had chronic impairment after a CVA at least one year prior to testing. Patients were aged between 40 and 78, with a mean age of 62 years (SD = 10.2). The patients were selected to show semantic impairment

(Table 8). Seven cases met the definition of semantic aphasia used by Jefferies & Lambon Ralph (2006) – i.e., multimodal deficits affecting comprehension tasks employing both words and pictures. The other cases were more mildly impaired; nonetheless, all cases were impaired at a demanding synonym judgement test, and they all had difficulty comprehending ambiguous words with multiple meanings, particularly following a miscue that primed the irrelevant interpretation of the word (Noonan et al., 2010). This suggests the patients' deficit was in controlling semantic retrieval such that the information being accessed was appropriate to the demands of the task: performance was relatively preserved when the task demands aligned with dominant aspects of knowledge, but were more impaired when the task required more unusual aspects of knowledge to be brought to the fore.

Older controls: Fifteen older adults, with a mean age of 73 years (SD = 8.1; range 55-84 years) were selected from a participant database at the University of York (ten female, five male). They were selected to provide an age-matched control group for our patient sample below. Participants had no prior history of brain injury, and showed unimpaired cognitive functioning on the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975), see Table 8 for their average performance on background semantic and executive assessments. They left school or college aged 18 (SD = 3.9 years).

-Table 8 about here-

Task and Design

This experiment used a thematic category matching task (as in Experiments 1 and 5), i.e., categories such as: PICNIC or HOSPITAL. Participants were asked to classify spoken words in terms of whether they were associated with these categories or not.

Materials: Stimuli were taken from Experiment 1. Twenty different thematic category labels were used (such as PICNIC) with 60 items in each category. 20 items were related to the category, including 10 targets that were strongly related to the category label, such as "*sandwich*"; and 10 that were distantly related, such as "*wasp"*, while the remaining 40 items were unrelated to the category (e.g. "*exam*") – these were recycled items from other categories. (See Appendix A in the Supplemental Material for a complete list of categories and items used).

Procedure: The experiment was presented using E-Prime 2.0 (Psychology Software Tools, Sharpsburg, PA). Testing was completed across two sessions. Category names were presented as written words that remained visible throughout the block to reduce demands on working memory. Participants were asked to press a button each time they heard a word that was related to the category, and not to press for unrelated words. The session began with 20 minutes of PSST testing (i.e. 5 categories out of the 20) at a presentation speed of 2 seconds. This was followed by 15 minutes of neuropsychological testing, followed by another 20 minutes of PSST testing (i.e., next 5 categories). The remaining 10 categories were presented in the second session, which followed the same procedure.

Results

Our main dependent measure was response sensitivity, since we aimed to evaluate the effect of semantic aphasia on the ability to detect targets belonging to each category, by the deadline imposed by the task. As before, we used (generalised) linear models (GLMs) for all analyses and entered RT as a covariate (i.e., the average speed for correct responses per condition per participant) in the analyses examining response sensitivity.

We examined the effects of three within-subjects factors in a GLM: (1) *relatedness* (targets with a strong or weak association with the category), (2) *set* (comparison of task performance in the first half of each session compared with the second half of each session, to assess the possibility of general cognitive fatigue); and (3) *within-category* position (comparison of task performance in the first half compared with the second half of each category), plus (4) *group* as a between-subjects factor. These predictors were entered in a fully-factorial model, including RT as a covariate. We also modelled *category number* (i.e., performance for each of the five categories within each set) and *session number* (i.e., day 1 or day 2 of testing) without interaction terms, to capture these aspects of the design. The model elicited a significant main effect of relatedness (see Table 9 for all *Wald* χ^2 and *p* values), and a significant two-way interaction between *group* and *relatedness* (see Figure 9): sensitivity was lower for weak items in comparison to strong items across both groups but separate analyses split by group showed that relatedness had a larger effect on performance for the patients: *Wald* χ^2 (1) = 101.43, $p < .001$, than for the controls: *Wald* χ^2 (1) = 34.64, $p < .001$. The patients showed poorer performance, relative to controls, for weak associations.

-Figure 9 about here-

There was no main effect of within-category performance, there was however a significant interaction between *relatedness* and *within-category* performance: there was greater within-category decline in performance for the weak targets and a subtle improvement in performance for the strong items towards the end of each category (see Figure 9). This interaction was significant in both groups; Controls: *Wald* χ^2 (1) = 18.98, $p < .001$; Patients: *Wald* χ^2 (1) = 8.718, $p = .003$, and contrasted with the pattern found for healthy young volunteers across multiple experiments above.

-Table 9 about here-

Summary of Experiment 6

The results showed that patients with semantic aphasia (relative to age-matched controls) were less efficient at retrieving weak associations, relative to stronger ones: in this way, the patients showed the expected pattern of poorer performance on trials requiring more controlled retrieval of semantic information. In addition, neither the patients nor the controls showed the within-category decline for strong associations seen repeatedly in healthy young participants. Since participants with deficits of semantic control did not show an increase in within-category decline in categorisation, our findings do not support the hypothesis that this decline in the PSST reflects an increasing need for controlled retrieval towards the end of each category – e.g., following the activation of a broad semantic field which would give rise to strong competition from semantically-related items. Instead, within-category decline in comprehension might be minimised in both patients with semantic aphasia and healthy older volunteers if this effect has its origins in the retrieval-induced suppression of semantically-related information on earlier trials. The flexible suppression of semantically-linked but non-target information might be an efficient way for younger participants to efficiently categorise targets, but this approach would be unhelpful or impractical for older volunteers and patients with a reduced capacity to control semantic retrieval. We note that, given some subtle changes in the design of this experiment (i.e., categories were tested across two sessions), we were not able to directly assess the effects of age on performance. We cannot exclude the possibility that the experiment lacked the sensitivity to recover differences between controls and patients, but the results do suggest that patients with deficits of controlled retrieval are not highly vulnerable to effects of within-category decline on the PSST. This observation further suggests that declining comprehension in the PSST has different origins from refractory effects in cyclical word-picture matching, since patients with semantic aphasia have been shown in previous studies to have increasing impairment across cycles in which the same stimuli are presented repeatedly (Gardner et al., 2012; Jefferies et al., 2007; Thompson et al., 2015). Potential reasons for this apparent difference between paradigms are considered below.

Discussion

Retrieval-related declines in categorisation were first assessed in healthy young adults across five experiments. Results demonstrated a cumulative decline in semantic categorisation as a consequence of sustained semantic retrieval: participants' ability to detect targets belonging to a particular category deteriorated even in the absence of item repetition/satiation. This effect was not equivalent to time-on-task, and could not be explained as a general decline in sustained attention as a result of fatigue, since many categories were tested back-to-back over the course of the experiments and there was a release from this phenomenon at category boundaries, when targets were no longer related to recently-categorised targets. The effects were observed in two different paradigms: both a vigilance paradigm, in which participants attempted to detect targets that were less frequent than distracters, and also in a 2AFC paradigm, where target and distracter items within and outside of the target category were presented equally often. This confirmed that healthy young participants were less able to categorise accurately towards the end of each category, and were not simply changing their response criteria following more experience with each category. The effect was largest for targets *strongly* related to the category, supporting the suggestion that this phenomenon is linked to spreading activation or inhibition between related concepts. However, it did not require a global semantic relationship between the targets within the category: within-category decline was seen across taxonomic, thematic and individual feature-based classification, suggesting it may be a fairly ubiquitous consequence of sustained semantic retrieval, at least in circumstances such as those created by the PSST paradigm, where the focus of retrieval is pre-defined. The effect was also multimodal (extending to a paradigm in which semantically-related pictures and words were interleaved), confirming that it is conceptual and not lexical in origin. Finally, it was increased by conditions of divided attention, suggesting that the capacity to maintain semantic retrieval within a category can be increased through the allocation of executive control.

The second part of this study examined how semantic aphasia might influence performance on the continuous categorisation task. Patients with semantic aphasia showed deficits in controlled retrieval (i.e., they maintained close-to-normal performance for strong associations, compared to age-matched controls, but had additional difficulties identifying weakly-associated targets). Categorising weakly-related items is thought to require greater semantic control, since there are fewer overlapping features for such items, and dominant yet currently-irrelevant features potentially have to be supressed in order to allow weakly-instantiated knowledge to come to the fore (Noonan et al., 2009; Whitney, Kirk, O'Sullivan, Lambon Ralph, & Jefferies, 2010). For example, to recognise that WASP fits with the category PICNIC, it is necessary to selectively focus on the feeding behaviour of wasps, and inhibit other information (e.g., wasps are insects, that make nests from wood). Despite this evidence of deficient semantic control, the patients with semantic aphasia resembled age-matched controls in the effects of within-category decline (and in both groups, there was little evidence of declining comprehension over successive targets within a category, in contrast to younger adults). This preserved performance is consistent with the hypothesis that within-category decline did not reflect strengthening competition between activated items (since patients with semantic aphasia would be expected to have difficulty resolving such competition); instead, the within-category decline seen in healthy young participants may have reflected the suppression of non-target semantically-related information on earlier trials. Patients with semantic aphasia and older healthy volunteers were apparently less susceptible to this effect, perhaps because they were not suppressing the non-target information to the same extent as younger participants.

Our results add to a growing body of work showing that conceptual processing can become less efficient following the retrieval of semantically-related items. The within-category effect in younger adults resembled the less efficient semantic retrieval that is seen when a sequence of related items are categorised (e.g., Campanella & Shallice, 2011; Harvey & Schnur, 2015; Wei & Schnur, 2015) – however, it occurred *without* the massed repetition and habituation of individual items common across these previous studies. The paradigm resembled a continuous picture naming task since items related to a category were interspersed with other items (Belke & Stielow, 2013; Belke, 2013; Howard et al., 2006; Kleinman, 2013; Navarrete et al., 2010; Oppenheim et al., 2010; Runnqvist et al., 2012; Schnur, 2014); yet it involved comprehension and categorisation rather than requiring speech output. Previous work examining categorisation in continuous paradigms has found *facilitation* (Belke, 2013; Riley et al., 2015). However, we were able to elicit declining performance by adapting this paradigm to require goal-driven categorisation, where the goal changed across successive blocks. We speculate below that this requirement to selectively focus retrieval on specific aspects of knowledge might have been critical to the effects we observed. The within-category decline in categorisation was also stronger when rapid processing was required. This resembles Campanella and Shallice's (2011) finding of deteriorating comprehension for repeated items in semantically-related sets when healthy participants had to respond by a deadline. Sensitivity to the speed of response suggests that a process that takes time (such the application of control to boost previously supressed representations) is essential to avoiding the detrimental effects of continuous categorisation.

There has been considerable discussion of: (i) whether cumulative interference occurs in comprehension paradigms (or whether these effects are restricted to picture naming); (ii) whether such effects arise at the level of lexical or semantic representations and (iii) whether these effects are short-lived (reflecting interference from on-going activation of semantically-related items), or are longer-lasting (reflecting weight-changes or adaptation within the underlying representations). Our data speak to all of these issues. First, within the literature on semantic access impairment, there has been debate about whether declining comprehension is restricted to auditory-verbal materials (Crutch & Warrington, 2008; Warrington & McCarthy, 1983; Warrington & Crutch, 2004), or whether it extends to non-verbal tasks (Forde & Humphreys, 1997; Gardner et al., 2012; Thompson et al., 2015). Similarly, opposing psycholinguistic studies have argued that semantic interference effects reflect lexical processes (e.g., Damian et al., 2001) or alternatively conceptual processes that extend to picture matching tasks (Wei & Schnur, 2015). We reasoned that if within-category decline effects emerge from within modality-specific representations, this pattern should be weaker in an 'interleaved' condition involving both words and pictures, as there would be more time for recovery between successive items, and/or fewer related targets presented within a modality to produce a decline in performance. In contrast, if these effects arise at a multi-modal conceptual level, they should be strong even when inputs of different modalities are interleaved. We observed a within-category decline in comprehension for interleaved words and pictures, suggesting that our results are unlikely to reflect either effects at a lexical-level or effects within the mappings between concepts and specific inputs (e.g., auditory word forms; structural descriptions of objects). We conclude this phenomenon originated within conceptual representations that are not specific to a particular input modality (as envisaged by Patterson, Nestor, & Rogers, 2007; Rogers et al., 2004; see also Wei & Schnur, 2015).

The semantic locus of the effect is also supported by our observation that the within-category decline in comprehension was most marked for strong associates of the target category: this suggests that spreading activation or inhibition between semantically-related concepts is a necessary condition for deteriorating categorisation. This finding, revealed across multiple experiments in healthy young volunteers, also speaks to two different mechanisms that have been put forward to explain declining comprehension. First, the literature on cyclical word-picture matching and semantic 'access' deficits has associated deteriorating performance with increasing competition within a set of items that are repeatedly presented (Gardner et al., 2012; Jefferies et al., 2007; Schnur, Schwartz, Brecher, & Hodgson, 2005; Thompson, Henshall, & Jefferies, 2016). By this view (and in contrast to our results), we might expect weak category members to show deteriorating performance under increasing levels of competition, since weak targets are less able to withstand strong competition. An alternative perspective is provided by the literature on continuous naming, where it has been suggested that semantically-related information is suppressed to facilitate target retrieval, and this effect makes semantically-associated items harder to retrieve on subsequent trials (Oppenheim et al., 2010). By this view, strong associates of the category should be suppressed to a greater extent than weak associates: for example, in the category PICNIC, suppression of a dominant category member such as FOOD might make it easier to retrieve other items that share few features with this item, such as SUNSHINE, and RUG. Our findings are consistent with this second account. In this way, the decline in categorisation that follows sustained semantic retrieval can be related to the phenomenon of retrieval-induced forgetting (see Anderson, 2003 for a review). This work shows that repeatedly retrieving one associate (e.g., FRUIT-ORANGE) renders a non-practised related associate (FRUIT-BANANA) less accessible. It is suggested that suppressing closely-related items facilitates retrieval by reducing competition, but this has consequences for the accessibility of the competing memory later. Retrieval-induced forgetting also affects strong exemplars of categories more than weak exemplars (Anderson, Bjork, & Bjork, 1994). The current study demonstrated these effects in a continuous categorisation paradigm, when there were more immediate consequences of earlier retrieval within the task.

Another theoretical debate concerns the underlying neurobiological mechanisms that give rise to deteriorating performance. While some studies (particularly those employing cyclical paradigms in which the same set of items are repeatedly presented at fast speeds) have assumed that *ongoing activation* of a set of competitors gives rise to declining performance (Campanella & Shallice, 2011; Schnur et al., 2006), research examining both continuous picture naming and retrieval-induced forgetting suggests that longer-term *weight changes* underpin these effects. Continuous naming studies have manipulated the lag between semantically-linked items and have found relatively long-lasting changes in accessibility that are unlikely to be explained by ongoing activation of conceptual information (Howard et al., 2006; Oppenheim et al., 2010). We conducted a lag analysis and obtained a similar pattern: categorisation was poorer in the second half of the category, particularly when there were lags of between one and three intervening items between successive targets. The experiment was not designed in a way that allowed us to examine longer lags, but targets that immediately followed a previous target without any intervening items did not show an effect of within-category decline, even though ongoing activation would have been at its strongest.

The findings from patients with semantic aphasia and age-matched controls are also consistent with the view that within-category decline in healthy young adults followed from the inhibition of aspects of meaning that were irrelevant or unhelpful during the classification of specific targets earlier in the list. In the PSST, dominant concepts for the category goal (e.g., SANDWICH for PICNIC) are likely to be quite distinct from the target word (e.g., RUG) and this goal-driven aspect of the paradigm might encourage healthy young participants to inhibit semantically-related yet currently-irrelevant information. While the inhibition of semantically-related items may have facilitated the retrieval of individual targets in student volunteers and given rise to the pattern of declining categorisation we have documented, people with a reduced capacity to control semantic retrieval (as a consequence of semantic aphasia or older age) should be less able to adopt this approach, eliminating or reducing declines in continuous categorisation in line with our findings.

Although we did not directly compare older and younger adults, recent work found that elderly participants show positive priming effects from previously-presented distractors that younger adults have inhibited (Amer & Hasher, 2014; Biss, Ngo, Hasher, Campbell, & Rowe, 2013). Advancing age is associated with poorer cognitive control, yet people continue to acquire factual information over their lifetime and tests of semantic knowledge show relatively little age-related decline (Burke & Mackay, 1997). Older adults may retrieve this semantic information in a different way, although further research is needed to fully document the effects of age and semantic control deficits on the PSST.

Our findings in patients with semantic aphasia further suggest that declining comprehension in the PSST has different origins from refractory effects in cyclical word-picture matching, since these patients have been shown by previous studies to have increasing impairment across cycles in which the same stimuli are presented repeatedly (Gardner et al., 2012; Jefferies et al., 2007; Thompson et al., 2015). These differences might be explained by the specific demands of the two tasks. As noted in the Introduction, the PSST requires participants to decide if each item fits within a specific category, which repeatedly changes over the course of the experiment – in this sense the task is more goal-driven than either naming or word-picture matching. In contrast, cyclical matching paradigms are stimulus-driven, but their structure might result in strong levels of competition, since they involve selecting items, supressing them when the target changes, and then re-selecting them: deficits of semantic control in patients with semantic aphasia following damage to left inferior frontal gyrus are argued to produce increased difficultly in target selection in the face of this growing competition. In contrast, when the target for categorisation continuously changes in the PSST, competition from previous targets is minimised, and effects of the suppression of related non-target items on earlier trials might be increased. Patients may show an overall deficit on the PSST because they are unable to focus retrieval on the current category goal, but as a consequence they do not show deteriorating performance. Further research is needed to directly manipulate the extent to which continuous categorisation is goal-driven, to establish if this factor is critical in creating the conditions for healthy young adults to show declining performance, and for patients with semantic access deficits to be relatively immune to such declines in performance.

Evidence across the experiments in healthy young participants also suggests that within-category decline is ameliorated through the application of cognitive control, which can overcome the earlier suppression of items that have now become targets. In young adults, the effect of within-category decline was larger towards the end of more executively-demanding paradigms, such as the specific feature-matching judgements in Experiment 3, perhaps because participants were no longer able or willing to constrain semantic retrieval in an effortful way to meet the demands of the task.

Similarly, within-category decline was greater in the interleaved condition of Experiment 4, which required greater attentional control to switch between words and pictures. Experiment 5 directly manipulated the availability of executive resources through the use of a secondary task to divide attention, and this appeared to increase the magnitude of the within-category decline in categorisation for highly-related items. Thus, it might be that the most substantial effects of within-category decline in goal-driven categorisation occur when: (i) new targets are highly related to the specified goal for categorisation and are therefore associated with previous targets; (ii) participants are young adults who use retrieval-induced suppression to facilitate efficient target retrieval and (iii) executive resources that could be used to control the negative effect of this previous retrieval are weak. In this way, our findings are readily related to contemporary accounts of semantic processing which envisage that heteromodal conceptual knowledge interacts with control processes to support context- and task-appropriate semantic retrieval (Jefferies, 2013; Lambon Ralph et al., 2016).

Our work shows that controlled retrieval of information pertaining to a specific topic is hard to sustain over time and increasing executive resources are required to overcome the negative effects of controlled retrieval on earlier trials. The phenomenon we have described would tend to promote an evolving pattern of retrieval that does not stay focussed on the same tight category but changes over time. Interestingly, this phenomenon is seen commonly in patterns of semantic retrieval that occur in everyday life, such as during mind-wandering and when conversations spontaneously shift topic (e.g., Humphries, Binder, Medler, & Liebenthal, 2006). In the real world, it might only be a matter of time before one of these alternative avenues becomes the focus of our retrieval.

Supplemental Material

The Supplemental Material can be found at the address (online address to be filled in)

References

Abdel Rahman, R., & Melinger, A. (2011). The dynamic microstructure of speech production: semantic interference built on the fly. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*(1), 149–61. https://doi.org/10.1037/a0021208

Almaghyuli, A., Thompson, H., Lambon Ralph, M. A., & Jefferies, E. (2012). Deficits of semantic control produce absent or reverse frequency effects in comprehension: Evidence from neuropsychology and dual task methodology. *Neuropsychologia*, *50*(8), 1968–1979. https://doi.org/10.1016/j.neuropsychologia.2012.04.022

Amer, T., & Hasher, L. (2014). Conceptual processing of distractors by older but not younger adults. *Psychological Science*, *25*(12), 2252–2258. https://doi.org/10.1177/0956797614555725

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language* (Vol. 49). https://doi.org/10.1016/j.jml.2003.08.006

Anderson, M. C., Bjork, E. L., Bjork, R., & Jordan, D. S. (2000). Retrieval-induced forgetting- Evidence for a recall-specific mechanism. *Psychonomic Bulletin & Review*, *7*(3), 522–530. https://doi.org/10.3758/BF03214366

Anderson, M. C., Bjork, R. A, & Bjork, E. L. (1994). Remembering can cause Forgetting: Retrieval Dynamics in Long-Term Memory. *Journal of Experimental Psychology*, *20*(5), 1063–1087.

Badre, D., Poldrack, R. A, Paré-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*, *47*(6), 907–18. https://doi.org/10.1016/j.neuron.2005.07.023

Badre, D., & Wagner, A. D. (2002). Semantic Retrieval, Mnemonic Control and Prefrontal Cortex. *Behavioral and Congitive Neuroscience Reviews*, *1*(3), 206–218.

Balota, D. A, & Black, S. (1997). Semantic satiation in healthy young and older adults. *Memory & Cognition*, *25*(2), 190–202. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9099071

Belke, E. (2008). Effects of working memory load on lexical-semantic encoding in language production. *Psychonomic Bulletin & Review*, *15*(2), 357–363. https://doi.org/10.3758/PBR.15.2.357

Belke, E. (2013). Long-lasting inhibitory semantic context effects on object naming are necessarily conceptually mediated: Implications for models of lexical-semantic encoding. *Journal of*

Memory and Language, *69*(3), 228–256. https://doi.org/10.1016/j.jml.2013.05.008

Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *The Quarterly Journal of Experimental Psychology*, (February 2013), 37–41.

Belke, E., & Stielow, A. (2013). Cumulative and non-cumulative semantic interference in object naming: evidence from blocked and continuous manipulations of semantic context. *Quarterly Journal of Experimental Psychology*, *66*(11), 2135–60. https://doi.org/10.1080/17470218.2013.775318

Biss, R. K., Ngo, K. W. J., Hasher, L., Campbell, K. L., & Rowe, G. (2013). Distraction can reduce age-related forgetting. *Psychological Science*, *24*(4), 448–455. https://doi.org/10.1177/0956797612457386

Bozeat, S., Lambon Ralph, M. a, Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, *38*(9), 1207–15. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10865096

Burgess, P. W., & Shallice, T. (1997). *The hayling and brixton tests. Bury St Edmunds: Thames Valley Test Company.*

Burke, D. M., & Mackay, D. G. (1997). Memory, language, and ageing. *Philos Trans R Soc Lond B Biol Sci*, *352*(1363), 1845–1856. https://doi.org/10.1098/rstb.1997.0170

Campanella, F., Mondani, M., Skrap, M., & Shallice, T. (2009). Semantic access dysphasia resulting from left temporal lobe tumours. *Brain*, *132*(1), 87–102. https://doi.org/10.1093/brain/awn302

Campanella, F., & Shallice, T. (2011). Refractoriness and the healthy brain: A behavioural study on semantic access. *Cognition*, *118*(3), 417–431. https://doi.org/10.1016/j.cognition.2010.08.005

Campbell, Hasher, L., & Thomas, R. (2010). Hyper-Binding: A Unique Age Effect, *294*(5551), 2536–2539.

Cipolotti, L., & Warrington. (1996). Does recognizing orally spelled words depend on reading? An investigation into a case of better written than oral spelling. *Neuropsychologia*, *34*(5), 427–440. https://doi.org/10.1016/0028-3932(95)00126-3

Cohene, L. S., Smith, M. C., & Klein, D. (1978). Semantic satiation revisited with a lexical decision task. *Memory & Cognition*, *6*(2), 131–140. https://doi.org/10.3758/BF03197438

Corbett, F., Jefferies, E., & Lambon Ralph, M. A. (2011). Deregulated semantic cognition follows prefrontal and temporo-parietal damage: evidence from the impact of task constraint on nonverbal object use. *Journal of Cognitive Neuroscience*, *23*(5), 1125–35. https://doi.org/10.1162/jocn.2010.21539

Crutch, S. J., & Warrington. (2003). Spatial coding of semantic information: Knowledge of country and city names depends on their geographical proximity. *Brain*, *126*(8), 1821–1829. https://doi.org/10.1093/brain/awg187

Crutch, S. J., & Warrington. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain*, *128*, 615–627. https://doi.org/10.1093/brain/awh349

Crutch, S. J., & Warrington, E. K. (2008). Contrasting patterns of comprehension for superordinate, basic-level, and subordinate names in semantic dementia and aphasic stroke patients. *Cognitive Neuropsychology*, *25*(4), 582–600. https://doi.org/10.1080/02643290701862290

Damian, Vigliocco, G., & Levelt, W. J. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, *81*(3), B77-86. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11483172

Davey, J., Thompson, H. E., Hallam, G., Karapanagiotidis, T., Murphy, C., De Caso, I., … Jefferies, E. (2016). Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. *NeuroImage*, *137*, 165–177. https://doi.org/10.1016/j.neuroimage.2016.05.051

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Forde, & Humphreys, G. W. (1995). Refractory semantics in global aphasia : On semantic organisation and the Access – Storage distinction in neuropsychology Refractory Semantics in Global Aphasia : On Semantic Organisation and the AccessStorage Distinction in Neuropsychology. *Memory*, (February 2013), 37–41.

Forde, & Humphreys, G. W. (1997). A Semantic Locus for Refractory Behaviour: Implications for Access Storage Distinctions and the Nature of Semantic Memory. *Cognitive Neuropsychology*, *14*(3), 367–402. https://doi.org/10.1080/026432997381529

Forde, & Humphreys, G. W. (2007). Contrasting effects of repetition across tasks: implications for understanding the nature of refractory behavior and models of semantic memory. *Cognitive, Affective & Behavioral Neuroscience*, *7*(3), 198–211. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17993206

Frenck-Mestre, C., Besson, M., & Pynte, J. (1997). Finding the locus of semantic satiation: an electrophysiological attempt. *Brain and Language*, *57*(3), 406–22. https://doi.org/10.1006/brln.1997.1756

Gardner, H. E., Lambon Ralph, M. A., Dodds, N., Jones, T., Ehsan, S., & Jefferies, E. (2012). The differential contributions of pFC and temporo-parietal cortex to multimodal semantic control: exploring refractory effects in semantic aphasia. *Journal of Cognitive Neuroscience*, *24*(4), 778–93. https://doi.org/10.1162/jocn_a_00184

Goodglass, H., & Kaplan, E. (1983). *The Boston Diagnostic Aphasia Examination.* Philadelphia: Lea & Febiger.

Gotts, S. J., & Plaut, D. C. (2002). The impact of synaptic depression following brain damage: a connectionist account of “access/refractory” and “degraded-store” semantic impairments. *Cognitive, Affective & Behavioral Neuroscience*, *2*(3), 187–213. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12775185

Harvey, D. Y., & Schnur, T. T. (2015). Distinct loci of lexical and semantic access deficits in aphasia: Evidence from voxel-based lesion-symptom mapping and diffusion tensor imaging. *Cortex*, *67*(March), 37–58. https://doi.org/10.1016/j.cortex.2015.03.004

Harvey, D. Y., & Schnur, T. T. (2016). Different loci of semantic interference in picture naming vs. word-picture matching tasks. *Frontiers in Psychology*, *7*(MAY), 1–19. https://doi.org/10.3389/fpsyg.2016.00710

Howard, Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: experimental and computational studies. *Cognition*, *100*(3), 464–82. https://doi.org/10.1016/j.cognition.2005.02.006

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *J Cogn Neurosci*, *18*(4), 665–679. https://doi.org/10.1162/jocn.2006.18.4.665

Jakobovits, L. A, & Lambert, W. E. (1962). Semantic satiation in an addition task. *Canadian Journal of Psychology*, *16*(2), 112–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14450948

Jefferies, E. (2013). The neural basis of semantic cognition: Converging evidence from neuropsychology, neuroimaging and TMS. *Cortex*, *49*(3), 611–625. https://doi.org/10.1016/j.cortex.2012.10.008

Jefferies, E., Baker, S. S., Doran, M., & Lambon Ralph, M. A. (2007). Refractory effects in stroke aphasia: a consequence of poor semantic control. *Neuropsychologia*, *45*(5), 1065–1079. https://doi.org/10.1016/j.neuropsychologia.2006.09.009

Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: A case-series comparison. *Brain*, *129*(8), 2132–2147. https://doi.org/10.1093/brain/awl153

Kay, J., Lesser, R., & Coltheart, M. (1992). *Psycholinguistic assessments of language processing in aphasia (PALPA).* Hove, UK: Lawrence Erlbaum Associates.

Kiss, G. R., Armstrong, C. A., & Milroy, R. (1973). An associative thesaurus of English. *Medical Research Council, Speech and Communication Unit, University of Edinburgh, Scotland.*

Kleinman, D. (2013). Resolving semantic interference during word production requires central attention. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(6), 1860–77. https://doi.org/10.1037/a0033095

Kroll, J., & Stewart, E. (1994). Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections between Bilingual Memory Representations. *Journal of Memory and Language*, *33*, *149–17*.

Lambert, W. E., & Jakobovits, L. A. (1960). Verbal satiation and changes in the intensity of meaning. *Journal of Experimental Psychology*, *60*(6), 376–83. https://doi.org/10.1037/h0045624

Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2016). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*. https://doi.org/10.1038/nrn.2016.150

Mirman, D. (2011). Effects of near and distant semantic neighbors on word production. *Cognitive, Affective & Behavioral Neuroscience*, *11*(1), 32–43. https://doi.org/10.3758/s13415-010-0009-7

Navarrete, E., Mahon, B. Z., & Caramazza, A. (2010). The cumulative semantic cost does not reflect lexical selection by competition. *Acta Psychologica*, *134*(3), 279–89. https://doi.org/10.1016/j.actpsy.2010.02.009

Neely, J. H. (1977a). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology. General*, *106*(3), 226–254. https://doi.org/10.1037/0096-3445.106.3.226

Neely, J. H. (1977b). The effects of visual and verbal satiation on a lexical decision task. *American*

Journal of Psychology, *90*(3), 447–459.

Noonan, K. A., Jefferies, E., Visser, M., & Ralph, M. A. L. (2013). Going beyond Inferior Prefrontal Involvement in Semantic Control : Evidence for the Additional Contribution of Dorsal Angular Gyrus and Posterior Middle Temporal Cortex. *Journal of Cognitive Neuroscience*, 1824–1850. https://doi.org/10.1162/jocn

Noonan, Jefferies, E., Corbett, F., & Lambon Ralph, M. a. (2009). Elucidating the nature of deregulated semantic cognition in semantic aphasia: evidence for the roles of prefrontal and temporo-parietal cortices. *Journal of Cognitive Neuroscience*, *22*(7), 1597–613. https://doi.org/10.1162/jocn.2009.21289

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews. Neuroscience*, *8*(12), 976–87. https://doi.org/10.1038/nrn2277

Pynte, J. (1991). The locus of semantic satiation in category membership decision and acceptability judgment. *Journal of Psycholinguistic Research*, *20*(4), 315–335. https://doi.org/10.1007/BF01074284

Radeau, M., Besson, M., Fonteneau, E., & Castro, S. L. (1998). Semantic, repetition and rime priming between spoken words: Behavioral and electrophysiological evidence. *Biological Psychology*, *48*(2), 183–204. https://doi.org/10.1016/S0301-0511(98)00012-X

Raven, J. . (1962). *Coloured Progressive Matrices Sets A, AB, B*.

Riley, E., McMahon, K. L., & de Zubicaray, G. (2015). Long-lasting semantic interference effects in object naming are not necessarily conceptually mediated. *Frontiers in Psychology*, *6*(MAY), 1–14. https://doi.org/10.3389/fpsyg.2015.00578

Rogers, T. T., Lambon Ralph, M. a, Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: a neuropsychological and computational investigation. *Psychological Review*, *111*(1), 205–35. https://doi.org/10.1037/0033-295X.111.1.205

Rowe, G., Valderrama, S., Hasher, L., & Lenartowicz, A. (2006). Attentional disregulation: a benefit for implicit memory. *Psychology and Aging*, *21*(4), 826.

Runnqvist, E., Strijkers, K., Alario, F. X., & Costa, A. (2012). Cumulative semantic interference is blind to language: Implications for models of bilingual speech production. *Journal of Memory and Language*, *66*(4), 850–869. https://doi.org/10.1016/j.jml.2012.02.007

Sambataro, F., Murty, V. P., Callicott, J. H., Tan, H. Y., Das, S., Weinberger, D. R., & Mattay, V. S. (2010). Age-related alterations in default mode network: Impact on working memory performance. *Neurobiology of Aging*, *31*(5), 839–852. https://doi.org/10.1016/j.neurobiolaging.2008.05.022

Schnur, T. T. (2014). The persistence of cumulative semantic interference during naming. *Journal of Memory and Language*, *75*, 27–44. https://doi.org/10.1016/j.jml.2014.04.006

Schnur, T. T., Schwartz, M. F., Kimberg, D. Y., Hirshorn, E., Coslett, H. B., & Thompson-Schill, S. L. (2009). Localizing interference during naming: Convergent neuroimaging and neuropsychological evidence for the function of broca's area. *Proceedings of the National Academy of Sciences*, *106*(1), 322–327. https://doi.org/10.1073/pnas.0805874106

Schnur, Schwartz, M., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language*, *54*(2), 199–227. https://doi.org/10.1016/j.jml.2005.10.002

Smith, L. C. (1984). Semantic satiation affects category membership decision time but not lexical priming. *Memory & Cognition*, *12*(5), 483–488. https://doi.org/10.3758/BF03198310

Smith, L., & Klein, R. (1990). Evidence for semantic satiation: Repeating a category slows subsequent semantic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(5), 852–861. https://doi.org/10.1037//0278-7393.16.5.852

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc*, *31*(1), 137–49. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10495845

Thompson-Schill, D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(26), 14792–7. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=25116&tool=pmcentrez&rendertype=abstract

Thompson, H. E., Henshall, L., & Jefferies, E. (2016). The role of the right hemisphere in semantic control: A case-series comparison of right and left hemisphere stroke. *Neuropsychologia*, *85*, 44–61. https://doi.org/10.1016/j.neuropsychologia.2016.02.030

Thompson, H. E., Robson, H., Lambon Ralph, M. A., & Jefferies, E. (2015). Varieties of semantic "access" deficit in Wernicke's aphasia and semantic aphasia. *Brain*, *138*(12), 3776–3792. https://doi.org/10.1093/brain/awv281

Tian, X., & Huber, D. E. (2010). Testing an associative account of semantic satiation. *Cognitive Psychology*, *60*(4), 267–90. https://doi.org/10.1016/j.cogpsych.2010.01.003

Vitkovitch, M., & Humphreys, G. W. (1991). Perseverant Responding in Speeded Naming of Pictures : It is in the Links. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *17*(4), 664–680.

Wagner, A. D., Desmond, J. E., Demb, J. B., Glover, G. H., & Gabrieli, J. D. (1997). Semantic repetition priming for verbal and pictorial knowledge: a functional MRI study of left inferior prefrontal cortex. *Journal of Cognitive Neuroscience*, *9*(6), 714–26. https://doi.org/10.1162/jocn.1997.9.6.714

Warrington, & Cipolotti, L. (1996). Word comprehension. The distinction between refractory and storage impairments. *Brain : A Journal of Neurology*, *119 (Pt 2*, 611–25. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8800952

Warrington, & Crutch, S. J. (2004). A circumscribed refractory access disorder: A verbal semantic impairment sparing visual semantics. *Cognitive Neuropsychology*, *21*(2), 299–315. https://doi.org/10.1080/02643290342000546

Warrington, & McCarthy, R. (1983). Category specific access dysphasia. *Brain : A Journal of Neurology*, *106 (Pt 4*, 859–78. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6652466

Wei, T., & Schnur, T. T. (2015). Long-Term Interference at the Semantic Level : Evidence From Blocked-Cyclic Picture Matching. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *42*(1), 149–157. https://doi.org/10.1037/xlm0000164

Wheeldon, L., & Monsell, S. (1994). Inhibition of Spoken Word Production by Priming a Semantic Competitor. *Journal of Memory and Language*, *33*, 332–356.

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M. A, & Jefferies, E. (2010). The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. *Cerebral Cortex (New York, N.Y. : 1991)*, *21*(5), 1066–75. https://doi.org/10.1093/cercor/bhq180

Figure Captions

Figure 1: Mean response sensitivity (*d'*) in Experiment 1 (Thematic-matching), for the first and second half of each category (within-category fatigue) and across the testing session (across-category fatigue), split by strong and weak targets, at the two presentation speeds. Error bars show *SE* of the mean.

Figure 2: Mean response sensitivity (*d'*) in Experiment 2 (Taxonomic-matching), for the first and second half of each category (within-category fatigue) and across the testing session (across-category fatigue). Error bars show *SE* of the mean.

Figure 3: Mean response sensitivity (*d'*) in Experiment 3 (Feature-matching), for the first and second half of each category (within-category fatigue) and across the testing session (across-category fatigue). Error bars show *SE* of the mean.

Figure 4: Mean response sensitivity (*d'*) in Experiment 4 (Effect across modalities), shown individually for the pictures and words modality in the interleaved and non-interleaved conditions, in the first and second half of each category (within-category decline), Error bars show *SE* of the mean.

Figure 5: Mean response sensitivity (*d'*) in Experiment 5 (Effect of divided attention), shown individually for the strong/weak targets, in the two conditions (single/dual), and split by first and second half of each category (within-category fatigue) and across the testing session (across-category fatigue). Error bars show *SE* of the mean.

Figure 6: Mean response sensitivity (*d'*), shown individually for the first and second half of each category (within-category fatigue) and across the testing session (across-category fatigue), for Experiments 1, 2, 3 and 5 (single condition), at the presentation speed of 1.1 seconds. Error bars show SE of the mean.

Figure 7: Average percentage of hits and false alarms for items within each category and across participants in Experiments 1 (Thematic-matching), 2 (Taxonomic-matching), 3 (Feature-matching) and 5 (Divided attention: single condition).

Figure 8: Mean response sensitivity (*d'*), shown individually for the first and second half of each category (within-category fatigue) and split by strong and weak trials, for Experiments 1, 4 and 5. Error bars show SE of the mean.

Figure 9: Mean response sensitivity (*d'*) for the first and second half of each category (within-category performance), split by strong and weak targets for patients and controls. Error bars show *SE* of the mean.

Table 1: Summary of significant results for response sensitivity from GLM and repeated-measures ANOVA analysis, examining effects of speed and relatedness, plus within-category and across-category changes in performance, in Experiment 1: Thematic-matching.

		Experiment 1: Thematic-matching	
		GLM (RT covariate)	*ANOVA*
Fixed effects:	***df***	***Wald χ2, p***	***F, p***
Across-category	(1, 23)	*p > .1*	*p > .1*
Within-category	(1, 23)	*p > .1*	*p > .1*
Relatedness	(1, 23)	***52.45, p < .001***	***50.26, p < .001***
Speed	(1, 23)	***40.25, p < .001***	***38.57, p < .001***
Interactions:			
Within-category x Relatedness	(1, 23)	***29.31, p < .001***	***28.09, p < .001***
Speed x Relatedness	(1, 23)	***17.62, p < .001***	***16.89, p < .001***
Speed x Within-category	(1, 23)	*p > .1*	*p > .1*
Across-category x Relatedness	(1, 23)	*3.24, p = .072*	*3.10, p = .091*

Footnote: Table presents two parallel analyses employing (i) mixed effects modelling (i.e., GLM preserving performance information for each category for each participant and treating participants as a random effect – this allowed RT per category to be included as a covariate of no interest) and (ii) analysis of variance. Other interaction terms were non-significant (p > .1).

Table 2: Summary of significant results for response sensitivity from GLM and repeated-measures ANOVA analysis, examining effects of across-category and within-category changes in performance, in Experiment 2: Taxonomic-matching.

		Experiment 2: Taxonomic-matching	
		GLM (RT covariate)	*ANOVA*
Fixed effects:	***df***	***Wald χ2, p***	***F, p***
Across-category	(1, 23)	*p > .1*	*p > .1*
Within-category	(1, 23)	***24.89, p < .001***	***23.85, p < .001***
Interactions:			
Across-category x Within-category	(1, 23)	*p > .1*	*p > .1*

Footnote: Table presents two parallel analyses employing (i) mixed effects modelling (i.e., GLM preserving performance information for each category for each participant and treating participants as a random effect – this allowed RT per category to be included as a covariate of no interest) and (ii) analysis of variance.

Table 3: Summary of significant results for response sensitivity from GLM and repeated-measures ANOVA analysis, examining effects of across-category and within-category changes in performance, in Experiment 3: Feature-matching.

		Experiment 3: Specific feature-matching	
		GLM (RT covariate)	*ANOVA*
Fixed effects:	***df***	***Wald χ2, p***	***F, p***
Across-category	(1, 23)	*p > .1*	*p > .1*
Within-category	(1, 23)	***13.43, p < .001***	***12.87, p = .001***
Interactions:			
Across-category x Within-category	(1, 23)	***6.18, p = .013***	***5.93, p = .025***

Footnote: Table presents two parallel analyses employing (i) mixed effects modelling (i.e., GLM preserving performance information for each category for each participant and treating participants as a random effect – this allowed RT per category to be included as a covariate of no interest) and (ii) analysis of variance.

Table 4: Summary of significant results for response sensitivity from GLM and repeated-measures ANOVA analysis, examining effects of modality and interleaving, plus within-category changes in performance, in Experiment 4: Cross-modality alternative-forced-choice decisions.

		Experiment 4: Across modalities	
		GLM (RT covariate)	*ANOVA*
Fixed effects:	***df***	***Wald χ2, p***	***F, p***
Within-category	(1, 21)	$p > .1$	$p > .1$
Modality	(1, 21)	$p > .1$	$p > .1$
Interleaved	(1, 21)	***15.72, p < .001***	***15.03, p = .001***
Interactions:			
Modality x Interleaved	(1, 21)	***7.39, p = .007***	***6.59, p = .018***
Interleaved x Within-category	(1, 21)	***4.85, p = .028***	***4.48, p = .046***
Modality x Within-category	(1, 21)	$p > .1$	$p > .1$
Modality x Interleaved x Within-category	(1, 21)	$p > .1$	$p > .1$

Footnote: Table presents two parallel analyses employing (i) mixed effects modelling (i.e., GLM preserving performance information for each category for each participant and treating participants as a random effect – this allowed RT per category to be included as a covariate of no interest) and (ii) analysis of variance.

Table 5: Summary of significant results for response sensitivity from GLM and repeated-measures ANOVA analysis, examining effects of condition (single/dual), relatedness, plus within-category changes in performance, in Experiment 5: Effect of divided attention.

		Experiment 5: Divided attention	
		GLM (RT covariate)	*ANOVA*
Fixed effects:	***df***	***Wald χ2, p***	***F, p***
Across-category	(1, 23)	*3.79, p = .052*	*2.52, p = .127*
Within-category	(1, 23)	*p > .1*	*p > .1*
Condition (single/dual)	(1, 23)	***6.81, p = .009***	***7.55, p = .012***
Relatedness	(1, 23)	***401.28, p < .001***	***327.25, p < .001***
Interactions:			
Relatedness x Within-category	(1, 23)	***6.60, p = .010***	***9.59, p = .005***
Condition x relatedness x within-category	(1, 23)	*3.55, p = .060*	*2.59, p = .123*

Footnote: Table presents two parallel analyses employing (i) mixed effects modelling (i.e., GLM preserving performance information for each category for each participant and treating participants as a random effect – this allowed RT per category to be included as a covariate of no interest) and (ii) analysis of variance. Other interaction terms were non-significant (p > .1).

Table 6: Summary of significant results from GLM and repeated-measures ANOVA analysis, examining across-category and within-category changes in performance across Experiments 1 (Thematic-matching), 2 (Taxonomic-matching), 3 (Feature-matching), and 5 (Effect of divided attention, single condition).

		Cross-Experiment comparison	
		GLM (RT covariate)	*ANOVA*
Fixed effects:	***df***	***Wald χ2, p***	***F, p***
Experiment	(1, 92)	***72.64, p < .001***	***27.79, p < .001***
Across-category	(1, 92)	*p > .1*	*p > .1*
Within-category	(1, 92)	***15.53, p < .001***	***14.83, p < .001***
Interactions (all n.s.):		*p > .1*	*p > .1*

Footnote: Table presents two parallel analyses employing (i) mixed effects modelling (i.e., GLM preserving performance information for each category for each participant and treating participants as a random effect – this allowed RT per category to be included as a covariate of no interest) and (ii) analysis of variance. Experiment was included as a between-subjects factor.

Table 7: Summary of significant results from GLM and repeated-measures ANOVA analysis, examining relatedness and within-category performance across Experiments 1 (Thematic-matching), 4 (Words modality), and 5 (Effect of divided attention).

		Cross-experiment relatedness comparison	
		GLM (RT covariate)	*ANOVA*
Fixed effects:	***df***	***Wald χ2, p***	***F, p***
Experiment	(1, 67)	***6.83, p = .033***	***4.38, p = .016***
Within-category	(1, 67)	***7.16, p = .007***	***7.09, p = .010***
Relatedness	(1, 67)	***289.14, p < .001***	***299.72, p < .001***
Interactions:			
Relatedness x Experiment	(1, 67)	***265.26, p < .001***	***45.72, p < .001***
Relatedness x Within-category	(1, 67)	***10.83, p = .001***	***11.12, p = .001***
Within-category x Experiment	(1, 67)	*p > .1*	*2.84, p = .066*

Footnote: Table presents two parallel analyses employing (i) mixed effects modelling (i.e., GLM preserving performance information for each category for each participant and treating participants as a random effect – this allowed RT per category to be included as a covariate of no interest) and (ii) analysis of variance. Experiment was included as a between-subjects factor. Other interaction terms were non-significant (p > .1).

Table 9: Summary of significant results from the GLM analysis for SA patients and age-matched controls – looking at the effects of group, relatedness, set and within-category performance, for our key dependent measures- response sensitivity, response accuracy and response times.

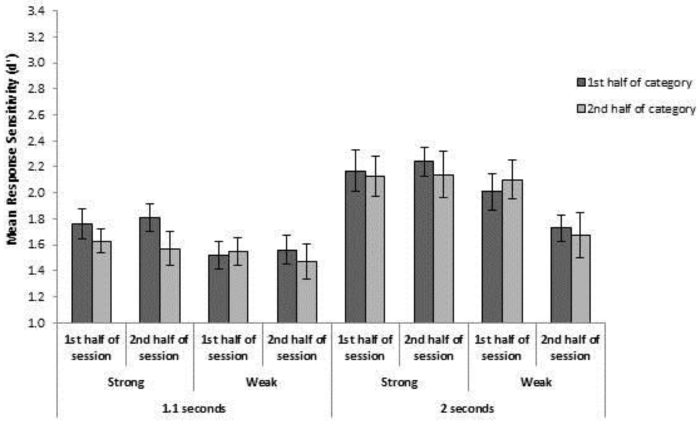
		SA patients vs. age-matched controls	
		GLM (RT covariate)	*ANOVA*
Fixed effects:	***df***	***Wald χ2, p***	***F, p***
Group	(1,23)	*3.19, p = .074*	***5.51, p = .027***
Relatedness	(1,23)	***126.95, p < .001***	***287.58, p < .001***
Set	(1,23)	*p > .1*	*p > .1*
Within-category	(1,23)	*p > .1*	*p > .1*
Interactions:			
Group x Relatedness	(1,23)	***8.08, p = .004***	*p > .1*
Group x Set	(1,23)	*3.42, p = .064*	*p > .1*
Relatedness x Within-category	(1,23)	***4.77, p = .029***	*p > .1*

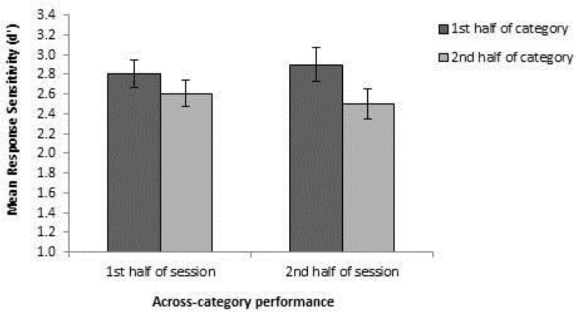
Footnote: Table presents analyses employing (i) mixed effects modelling for response sensitivity (i.e., GLM preserving performance information for each category for each participant and treating participants as a random effect – this allowed RT per category to be included as a covariate of no interest).

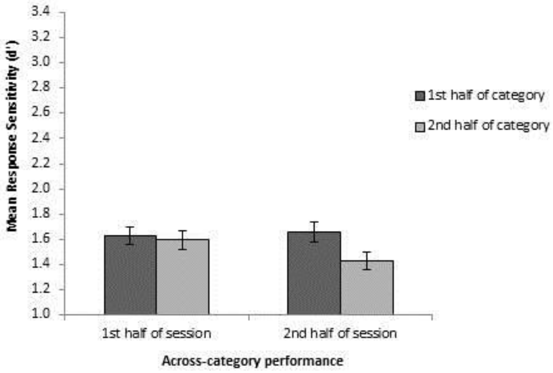
Table 8: Background neuropsychological data for each patient

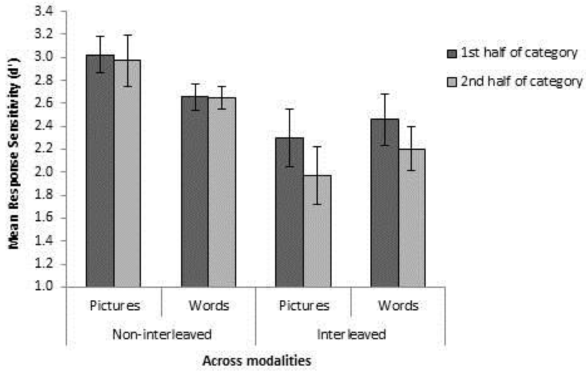
	Max score	Control mean	Cut-off	EKD	ONY	YHE	SSR	RTJ	NNZ	NHY	NGW	ESU	NNF	LHN
r tasks:														
	64	64	63	64	63	62*	52*	63	64	62*	64	62*	60*	62*
res	64	59	53	58	60	61	54	61	53	57	56	45*	45*	44*
ls	64	61	57	63	58	60	57	56*	61	52*	53*	59	29*	43*
Judgement	96	95	91	90*	87*	81*	87*	81*	78*	76*	74*	66*	71*	59*
e canonical	37	36	34	NA	36	37	33*	37	37	35	35	37	29*	31*
e non-canonical	37	34	29	NA	32	29	22*	32	26*	22*	21*	34	14*	13*
y: cues	60	60	59	NA	52*	54*	47*	57*	50*	51*	40*	43*	39*	35*
y: miscues	60	59	57	NA	50*	45*	39*	54*	42*	34*	22*	30*	27*	23*
e tasks:														
ing	23	23	17	23	23	22	23	21	19	5*	12*	1*	16*	23
	36	33	28	32	29	33	34	33	21*	30	24*	19*	31	29
	54	33	28	39	45	30	31	39	31	23*	26*	24*	18*	7*
ical deficits:														
ed WPM		NA		NA	58	37	0*	38	54	37	12	60	9	18
epetition	80	NA	73	NA	NA	77	1*	7*	74	79	75	78	42*	71

* Denotes impaired performance. NA = not available. Patients are arranged according to composite semantic severity scores; this is a single factor extracted from WPM = word picture matching, CCT = Camel and Cactus Task (both from Bozeat et al., 2000), and synonym judgement. RCPM = Raven's Coloured Progressive Matrices (Raven, 1962). BSRA = Brixton Spatial Attainment Task (Burgess & Shallice, 1997). PALPA = Psycholinguistic Assessments of Language Processing in Aphasia (Kay, Lesser, & Coltheart, 1992). Cookie theft description assesses fluency (words-per-minute; Goodglass & Kaplan, 1983)

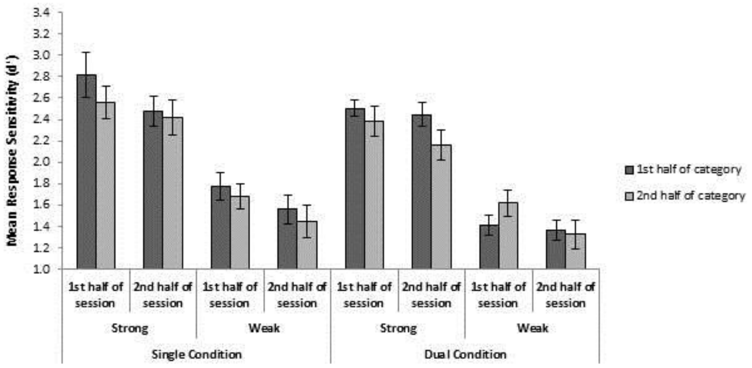

3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
Mean Response Sensitivity (d')
1st half of category
2nd half of category
1st half of session
2nd half of session
1st half of session
2nd half of session
1st half of session
2nd half of session
1st half of session
2nd half of session
Strong
Weak
Strong
Weak
1.1 seconds
2 seconds

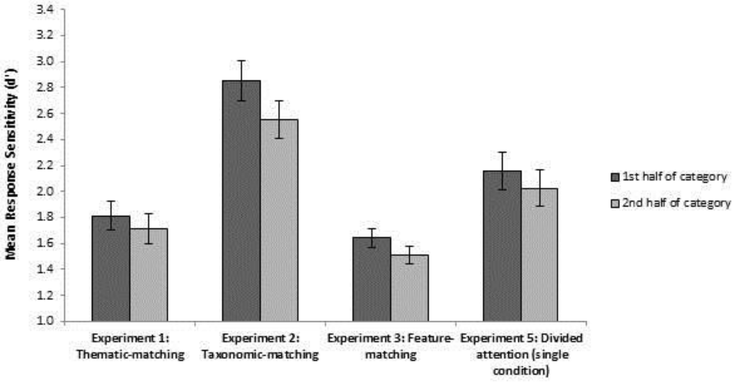

Mean Response Sensitivity (d')
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
1st half of category
2nd half of category
1st half of session
2nd half of session
Across-category performance


Mean Response Sensitivity (d')
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
1st half of category
2nd half of category
1st half of session
2nd half of session
Across-category performance

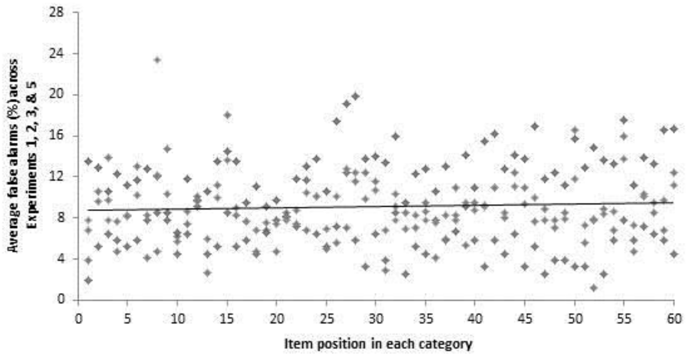


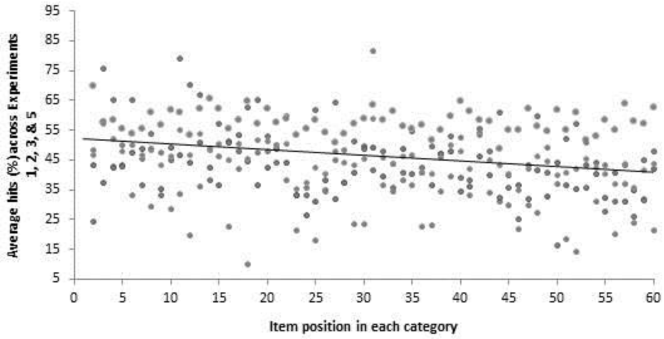
Mean Response Sensitivity (d')
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
1st half of category
2nd half of category
Pictures
Words
Pictures
Words
Non-interleaved
Interleaved
Across modalities

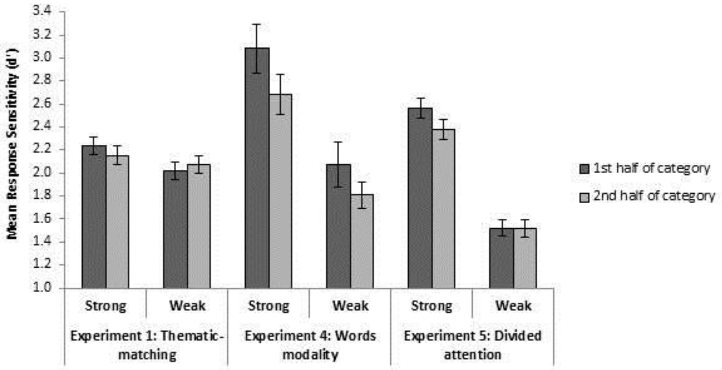

Mean Response Sensitivity (d')
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
1st half of session
2nd half of session
1st half of session
2nd half of session
1st half of session
2nd half of session
1st half of session
2nd half of session
Strong
Weak
Strong
Weak
Single Condition
Dual Condition
1st half of category
2nd half of category



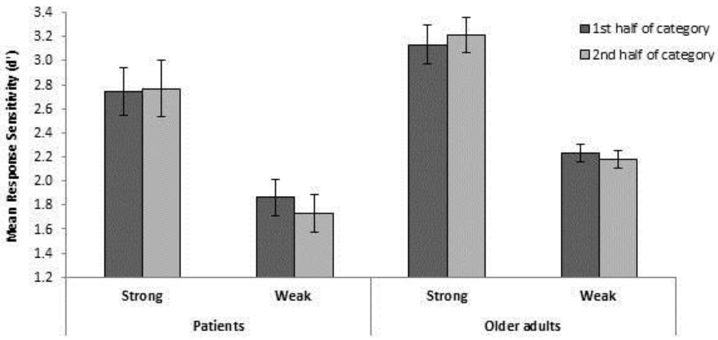
Mean Response Sensitivity (d')
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
Experiment 1: Thematic-matching
Experiment 2: Taxonomic-matching
Experiment 3: Feature-matching
Experiment 5: Divided attention (single condition)
1st half of category
2nd half of category


Average false alarms (%) across Experiments 1, 2, 3, & 5
Item position in each category
0
4
8
12
16
20
24
28
0
5
10
15
20
25
30
35
40
45
50
55
60


Average hits (%) across Experiments 1, 2, 3, & 5
95
85
75
65
55
45
35
25
15
5
0
5
10
15
20
25
30
35
40
45
50
55
60
Item position in each category



Mean Response Sensitivity (d')
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
Strong
Weak
Strong
Weak
Strong
Weak
Experiment 1: Thematic-matching
Experiment 4: Words modality
Experiment 5: Divided attention
1st half of category
2nd half of category


1st half of category
2nd half of category
Mean Response Sensitivity (d')
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
Strong
Weak
Patients
Strong
Weak
Older adults